Appendix:
Integrating econometric land use models with ecological modeling of ecosystem services to guide coastal management and planning

May 10, 2019

Contents

List of Tables

List of Figures

A Data sources and selection

This appendix identifies data sources, includes a brief overview of differences in local demographic and socioeconomic conditions among the three counties that supplements the discussion presented in section 3, describes in detail the steps followed to arrive at the final sample for estimation, evaluates the predictive power of the developed models, and discusses additional econometric considerations for modeling. Table A.1 specifies name and access for the sources consulted. Tables A.2–A.4 present county land use and zoning codes. Table A.5 shows vacant lands left out of the estimation sample by zoning or land use code. Figure A.1 shows a map of housing unit growth from 1990 to 2016 by census track in the three counties that conform study area. Figure A.2 shows the location of flooding events in the study area since 1996 in relation to the parcels available for analysis. Figure A.3 shows activity in the property markets in all three counties from 1950 to 2016.

A.1 Study area

For the empirical analysis, data pertains to three coastal counties in South Carolina that surround the Charleston metropolitan area: Charleston, Berkeley, and Georgetown. These counties are part of what is known as "the Lowcountry,"a region along the coast of South Carolina that is well known for its historic legacy, cultural importance, natural environment, and fast–growing economies.

The major forces driving increased flooding and increased damages from floods in the study area are urban and climatic pressures. This section provides a brief description of general socioeconomic trends in South Carolina followed by some observations on the economic outlook for each county in the study area. Also, the section contains a synthesis of historical flood and storm data for South Carolina, and describes recent trends in flooding pressures in the study area and how they compare to the occurrence of flooding events in South Carolina.

A.1.1 Local urbanization trends

The US Census identifies urban areas along coastal watershed counties in the Southeastern US to be among the fastest–growing areas in the country (US Census, 2015). Like other coastal areas in South Carolina, two of the counties being analyzed, have seen a rapid increase in urbanization in the last decade. Between 2010 and 2016, the urban population increased by 12.9% in Charleston and 17.8% in Berkeley, outpacing the nation's overall growth rate of 9.7 percent for the same period. However, in Georgetown, urban population growth was much slower, at 2.1%.[1]

Domestic migration and the influx of seasonal residents are important factors behind this large population growth. As such, changes in local property markets closely reflect the preferences of retirees, tourists, and visitors (NOAA, 2016). In recent years, there has been a rapid increase in development of infrastructure to support tourism and retirement communities, and much of that new development has occurred near or on the coast. In fact, according to the latest Census, coastal counties in South Carolina lead the national statistics on both absolute and relative increases in seasonal housing units among coastal counties. In 2010, South Carolina had the third largest share of seasonal housing units in the nation (12%), an important change since the 2000 Census when SC was not even among the top six states in regards to seasonal housing (NOAA, 2013).

Figure A.1 shows a map of housing unit growth from 1990 to 2016 by census track in the three counties that conform study area. Panel (a) in figure A.1 shows the number of structures built after 1990 recorded in the 2016 American Community Survey (ACS), and panel (b) shows the percentage growth in housing stock from 1990 to 2016. As shown by the figure, new development is concentrated around urban centers and near the coast, which is consistent with national trends.[2]

An examination of the most recent national land cover data indicates that much of the new development is dominated by medium and high intensity type of growth, although it is accompanied by growth that is more suburban in nature. Respectively, open space–, low–, medium–, and high–intensity development cover 3.2%, 2.4%, 1%, and 1.15% of the land in the study area.[3] Figure 3.4 also shows that the predominant land cover classes in the study area are woody wetlands, evergreen forests, and emergent herbaceous wetlands, which respectively cover 36%, 31%, and 13% of the study area. Pastures, herbaceous, and cultivated crops lands are concentrated towards the

[1]For reference, the nation's urban population increased by 12.1 percent from 2000 to 2010.

[2]NOAA identifies 672 coastal counties in the US, and although these counties make up more or less 25 percent of the country's area, they account for almost 46 percent of development. Recent data shows that between 2000 and 2010, around 1,880 building permits were issued per day in coastal counties while the total building permits issued in inland counties was around 2,160 (NOAA, 2013).

[3]Areas with mostly vegetation in the form of lawn grasses such as as local parks, golf courses, or large–lot single family parcels, are considered open space development. In these areas, less than 20% of the total cover is accounted for by impervious surfaces. Low intensity development areas are those most commonly used in single–family housing units, where impervious surfaces generally account for 20% to 47% of total cover. Medium intensity development areas, are most commonly single–family units where impervious surfaces account for 50% to 79% of the total cover. In turn, high intensity development areas are where people reside or work in high numbers and where impervious surfaces account for more than 70% of the total cover (e.g., apartment complexes).

northwest edge of the studied region, and together constitute 6% of the area of interest. The remaining 6% is covered by open water.[4]

Table 1 presents change in land cover composition in the three counties between 2001 and 2011 and makes clear that decreases in forest covers, pasture lands, and woody wetlands are likely giving way to the observed increase in medium– and high–intensity development land uses. Table 3.1 also shows that forest lands were consistently developed at larger rates than agricultural lands, and changes from agricultural land (pastures, hay, and crops) to forest were also important. In Charleston county, increases in area covered by developed surfaces are offset by decreases in forest and pasture covers. Interestingly, shifts from low–intensity to high–intensity development between 2006 and 2011 are also observed, as would be expected in exurban development. In Berkeley county, the large increase in medium– and high–intensity development (52% and 48%, respectively) is accompanied by a substantial reduction in forest (8.21%) and woody wetlands covers (1.45%). Finally, in Georgetown, there is an increase in all types of development but there are important changes in agricultural and commercial agricultural covers (pastures, hay, and crops), reflecting the rural character of the county.

A.1.2 History of floods in the study area

In South Carolina, over 320,000 residents live in flood risk areas, which are defined by the Federal Emergency Management Agency (FEMA) as locations that have a 1% or greater chance of being flooded in any given year (these areas are also referred to as subject to 100–year floods). Flooding is the second most common natural hazard in the state, affecting every county (NOAA, 2013).

In the past 20 years, floods in South Carolina and the study area have increased in frequency, and the costs associated with flood events have also increased.[5] Between 1996 and 2016 there were 1,628 flooding events in South Carolina (i.e., one every 26 hours).[6]

[4]The main inland water bodies are Lake Marion and Lake Moultrie, covering over 170,000 acres.

[5]The National Oceanic and Atmospheric Administration (NOAA) National Weather Service (NWS) division keeps a comprehensive database with records on flooding events from 1996 to present. Flooding events in NWS database include flash floods, floods, and coastal floods. Other types of storms recorded in the database include storm surges or tides, thunderstorms, hurricanes, "heavy rain"and tropical storms. The discussion below focuses only on flash floods, floods and coastal floods.

[6]NOAA defines a flooding event as a smaller geographic area that together aggregate to a specific storm or flood episode (i.e., one flooding episode can impact multiple locations, each of which would be a flooding event). In turn, a flooding episode is defined as any occurrence of storms and other significant weather phenomena having sufficient intensity to cause loss of life, injuries, significant property damage, and/or disruption to commerce. Between 1996 and 2016, there were 677 unique "flooding episodes"(i.e., one every 10.8 days) in South Carolina.

The number of flooding events with economic consequences consistently increased over time (NWS, 2018). In the study area, there were 436 flooding events between 1996 and 2016, which together summed up to almost $28 million in damages—representing 43% of overall property damages from all types of storms in the study area and over 15% of overall property damages from floods in South Carolina since 1996.[7] The storms data does not yet include information from damages caused by the 2017 Hurricane Irma. Noting that, the costliest year for flooding in the dataset was 2015 (see figure 3.6) when a large storm resulted in catastrophic flooding in South Carolina.[8] That year, damages in summed up to $24.75 million—around a fifth of the state's flooding costs.

Figure A.2 shows the location of flooding events in the study area since 1996 in relation to the parcels available for analysis. It is apparent from this map that there is heterogeneity across parcels in their exposure to flooding risks. Most flooding events occurred in Charleston county (over 75%). Also, damages caused by the average flooding event in Charleston were $4 to $5.25 million higher than in Berkeley and Georgetown.[9]

A.2 Overview of county differences

Charleston county is the largest county in SC by land and water area. The county has a total area of 1,358 square miles, 33% of which are covered in water, leaving approximately 916 square miles of land. It is also the third most populous county in the state with an estimated population of 389,262 people, or 8% of the state's population. In 2015, the median household income in Charleston is $57,079, just below the US median household income ($57,230 in 2015) but above the median household income in South Carolina ($47,238), and 7.07% higher than the previous year.

Charleston is also densely populated, with about 89% of the county's population being classified as urban, which is well above the state's share of urban residents (71% in the 2010 Census). In the

[7] The 1,628 flooding events experienced in South Carolina since 1996 caused almost $171 million losses in property damages, which is almost 25% of the overall cost from damages caused by all types of storms in the state. In the last 20 years, the state saw 5,898 storms. Of those, 5,815 caused property damages summing up to $687.5 million in damages. The average flood event caused $16.65 million in property damages and the average damage per affecting event was $248,300.

[8] It is worth mentioning that South Carolina has not been directly impacted by a hurricane since Hurricane Hugo in 1989. However, it is not uncommon for large storm systems over the southeaster states to draw moisture from hurricanes and subsequently result in flooding. In fact, South Carolina experienced property damages from floods every year since 1996.

[9] Analyses by NOAA show that regions within the study area that are the most vulnerable to changes in sea level rise are those around the Santee River (on the border between Berkeley and Charleston counties) and the Wadmalaw River (which has its delta 24 miles South of the City of Charleston).

2014 ACS, population density was 416 people per square mile, which is over 4.5 times the national average of 88.8 people per square mile. Between 2000 and 2015, the population in Charleston county grew by 25%, or at an annual rate of 8.8%, and the housing stock grew by almost 34,600 units, or 24.5%. Finally, Charleston county is the fastest growing county in the state and the county's seat, the City of Charleston, is one of the fastest growing metropolitan areas in the nation.

Berkeley county is adjacent to Charleston county (see figure 2) and has an area of 1,229 square miles, most of which are terrestrial (1,099 square miles). Seventy–one percent of the county's population is classified as urban, and with a population of 202,786, population density in the county is 180.4 people per square mile, making Berkeley county less than half as densely populated as Charleston county.

In 2015, the median household income of the 69,030 Berkeley households was $52,506, which is less than the median annual income in Charleston but represents a 1.3% increase from the previous year. Berkeley county also remains behind Charleston county in terms of absolute population and construction levels. However, in relative terms, Berkeley county has experienced faster population and development growth than Charleston county. Between 2000 and 2015, the population in Berkeley grew by 41.7% (i.e., 11.4% annually), and the housing stock grew by almost 40% (or 22,000 units).

Georgetown county, the northern–most county in the study area, has a total area of 1,035 square miles, of which 814 are terrestrial. Georgetown county is similar in area and physical characteristics to Charleston and Berkeley, but is very different in terms of socioeconomic conditions and economic history. Georgetown's population growth of approximately 1% per year is much slower than that of Charleston and Berkley (8.8% and 11.4%, respectively). With a population of 60,572 and a population density of 74.2 people per square mile, Georgetown is much more rural than both Berkeley and Charleston, and only 58.8% of the county was classified as urban in the 2010 Census.

The median household income in Georgetown in 2015 grew faster than in Berkeley, with a growth of 3.02%. Nevertheless, at $42,835, it remained well below the median income in the other two counties. Between 2000 and 2015, population grew much slower in Georgetown than in Charleston and Berkeley counties, while growth in the housing stock was half the growth in Berkeley but similar to that in Charleston (however, Charleston county is much more densely populated and therefore has smaller margins to expand in this regard). In this time period, the population in Georgetown grew by 9.4% (i.e., half as fast as in Charleston and even slower compared to Berkeley) and the housing stock grew by almost 20% (or 5,650 units).

Figure A.3 shows market trends and differences in sales volumes across counties. Panel (a) shows the number of market transactions per year in Charleston county, panel (b) shows historical sales in Berkeley county, and panel (c) presents the number of market transactions in Georgetown county. In general, market tendencies shown by the number of transactions of residential or agricultural parcels are mirrored across counties. However, much of the difference between counties is found in the volume of sales and recorded prices (prices are not shown in the figure). Overall, Charleston and Berkeley have more active markets than Georgetown.

As shown in figure A.3, there was limited activity in property markets in all three counties until late 1970's, when they started to rapidly pick up.[10] The early 2000's saw the end of that period of rising market activity. After a short slump, there was a second boom period where market activity increased more rapidly than before and in which the number of transactions overpassed the 2000 level, reaching a maximum number of transaction in 2005. There was a sharp decline around the time of the financial crisis with the minimum number of sales being recorded in 2008 for Charleston, 2012 for Berkeley, and 2009 in Georgetown. Sales began to recover in 2012, and in 2015, the number of transactions was already above 2005 levels.

A.3 Sample selection procedure

The empirical analysis is applied to a dataset with geographical, biophysical, climatic, and socioeconomic data for a subset of all properties in tax assessor databases. Considerable effort was devoted to arrive at the final parcel sample used in estimation as each dataset had to be carefully pre–processed to accurately characterize the use given to each parcel in the study area.

The first step in constructing the final sample of parcels, was to process the 2016 databases provided by each county's tax assessor office. Specifically, duplicated observations, observations with missing parcel identification numbers, and observations with recording errors were removed. Also, when necessary, separate available datasets provided by tax assessor offices were cleaned and merged to the basic dataset with transactions data to create a complete database with zoning information, building characteristics, year of development, and price information. To verify the ac-

[10]This could be because transactions were not recorded into tax assessors' databases.

curacy of the information in these datasets, complementary GIS parcel layers were used to estimate parcel size for each parcel in the tax assessor data. When appropriate, the area recorded in the tax assessor data was replaced by a more accurate measure of parcel size generated in GIS.

In a second stage, using zoning codes or property class codes, parcels were grouped into four classes: (1) undeveloped, (2) developed residential, (3) developed industrial, commercial, or mixed uses (which include residential properties under apartment complexes, condominiums, row houses, town houses, et.), and (4) public property. Parcels in the latter two groups were not used.[11]

Figure 4 in the main text shows the location of parcels in the sample relative to parcels that were not included and are categorized by land use. Figure 5 provides supplementary location information. Specifically, panel (a) of figure 5 presents the location of only the parcels included in the final sample and indicates whether they were considered developed or undeveloped, and panel (b) shows the location of parcels in the sample relative to land cover classes in the study area. As shown by the figure, the landscape pf parcels left out from the analysis is dominated by public lands, wetlands and marshes, and open water.

Several property class codes were grouped together to represent each of the land uses considered in the empirical analysis: residential uses and agricultural uses.[12] The sample of parcels in Charleston county that were selected as developed was restricted to parcels zoned as single family residential and residential development lots. In turn, parcels included in the undeveloped sample were those in vacant residential zones (as long as the parcel's reported building square footage was zero, otherwise, the parcel was considered developed), agricultural districts, and lots occupied by mobile homes. The reason why lots occupied by mobile homes were considered undeveloped is that these are easily removable structures, and are therefore seen as perfect substitutes for vacant lots. Also, to avoid grouping rural homes with undeveloped parcels, data on square footage of residential buildings was used to sort farms from rural homes. Hence, properties in agricultural zones but with residential buildings were considered developed.

[11] Residential development is generally considered the leading edge of development (Benson, 2009). Thus, industrial, commercial and mixed use properties are removed from the data. In addition, by keeping this restriction, the analysis is not subject to inconsistencies between zoning codes and *actual* uses of the land. These inconsistencies are rather common as shown by the tables in Appendix A that present zoning codes from the three counties.

[12] Zoning and property class codes represent the primary land use upon which the tax assessor bases its assessment of the property's value. To arrive at a satisfactory decision rule, county land use codes, zoning classifications, and parcel characteristics were carefully studied. At certain points, Google Earth was used to verify the accuracy of these codes.

To classify data from Berkeley county, an analogous procedure was followed to define the sample of developed residential and undeveloped parcels. Similar to Charleston county, only parcels zoned as single family residential, rural residential, agricultural districts, or lots occupied by mobile homes were considered for the study. To make sure no parcels zoned as commercial, industrial, mixed uses or having multifamily residential structures were included in this sample, the sample of developed parcels included properties with no assessed commercial value and no square footage of commercial buildings. An analogous filter was applied to data from Georgetown county: vacant lots, lots occupied by mobile homes, and agricultural properties with no residential buildings were grouped into the sample of undeveloped parcels. In turn, occupied single family residential properties, resort lots, and villas were considered developed. Publicly owned properties and properties with industrial, commercial, or mixed uses were excluded from both samples.

The third step in constructing the data for this study, was to drop observations from the developed sample with no information on key variables for the analysis, namely information regarding year of construction and basic information on building characteristics. The fourth step was to merge the parcel data with the spatial variables collected from other public sources and the proximity and buffering measures generated in ArcGIS.

The final step was to restrict the sample of parcels used in the predictive hedonics models according to the suggestions derived from a careful model selection process which followed machine learning validation techniques. The final selection for the land use change model had 53,228 undeveloped parcels and 129,634 developed parcels. In turn, the sample for the residential and agricultural hedonic models, which is composed by parcels which sold in recent years and had complete information, had 33,364 and 6,725 parcels, respectively.

A.3.0.1 Word of caution

Before moving on, it is important to address concerns over making errors in the process of separating developed from undeveloped parcels. A first concern is that of falsely classifying an undeveloped parcel in the sample as developed (i.e., a type I error).

In this analysis, type II errors (i.e., falsely labeling a developed parcel as undeveloped) are not problematic. If type II errors occur systematically, information about residential parcels will contaminate the sample of undeveloped parcels and result in biased predicted undeveloped values of developed parcels. Since development is assumed to be irreversible in this study, residential parcels are always predicted to develop, regardless of the corresponding predicted values. However, if undeveloped parcels are systematically falsely labeled as developed, the land use change model may mis predict development patterns in the study area. With type I errors, information of agricultural parcels would bias the hedonic analysis of residential parcels and make the resulting predicted developed values of undeveloped parcels more likely to reflect characteristics of undeveloped parcels than of developed parcels. If developed parcels are generally more valuable than undeveloped parcels, the prevalence of type I errors would cause the predicted developed values of undeveloped parcels to be biased downwards, making the predicted probabilities of development artificially smaller. Alternatively, if undeveloped parcels are generally more valuable than developed parcels, the predicted probabilities of development will be biased upwards and the land use change model will over predict development.

Another source of concern is the effect of not including relevant parcels in the sample. Table A.5 shows vacant lands left out of the estimation sample by zoning or land use code. The table suggests that the majority of vacant lands not considered in this analysis are public lands, parcels that are undevelopable, parcels without a current classification, or parcels that are classified as agricultural but may also serve as residential dwelling. Parcels for which particular spatial data was not available were also left out from the sample. An important number of those are found in the Northeast of Charleston, where data on soil quality was unsuitable for the analysis.

Table A.1: Data sources.

Layer	Source	Website
Land cover data	National Land Cover Database	www.mrlc.gov/finddata.php
Soil data	National Resources Conservation Service	www.nrcs.usda.gov/wps/portal/nrcs
Geographic names	United States Geological Survey	gdg.sc.egov.usda.gov
National hydrography dataset	United States Geological Survey	gdg.sc.egov.usda.gov
2015 Primary and secondary roads	United States Geological Survey	gdg.sc.egov.usda.gov
National Elevation dataset	United States Geological Survey	gdg.sc.egov.usda.gov
SC public access to beaches	SC Dept. of Health and Environ. Control	www.scdhec.gov
Socioeconomic data	American Community Survey	factfinder.census.gov
Housing Price Index	Federal Reserve Bank	fred.stlouisfed.org/series
Tax assessor data Charleston	Tax assessor office	www.charlestoncounty.org
Tax assessor data Berkeley	Tax assessor office	www.berkeleycountysc.gov
Tax assessor data Georgetown	Tax assessor office	www.georgetowncountysc.org
Charleston county GIS office	GIS office	www.charlestoncounty.org
Berkeley county GIS office	GIS office	gis.berkeleycountysc.gov
Georgetown county GIS office	GIS office	www.georgetowncountysc.org
Other consulted sources		
Wood Mills in South US 2015	Forest economics and policy research	www.srs.fs.usda.gov/econ/data/mills/
SC hurricane evacuation routes	SC Emergency Management Division	www.scdhec.gov/HomeAndEnvironment/
Layer	Source	Website
Storm database	NOAA's National Weather Service	www.ncdc.noaa.gov/stormevents/ftp.jsp
Building permits database	State Of the Cities Data Systems	socds.huduser.gov/permits/
SC department of transportation		info.scdot.org/sites/GIS/
Unites States National Conservation Easement Database		www.conservationeasement.us/
South Carolina Department of Natural Resources		www.dnr.sc.gov/
Flood hazard maps	Federal Emergency Management Agency	www.floodmaps.fema.gov/NFHL/
USGS Inundation maps	United States Geological Survey	water.usgs.gov/floods/events/2015/Joaquin/
CERA group storm simulations	Coastal Emergency Risk Assessment	nc-cera.renci.org/cgi-cera/cera-nc.cgi

Table A.2: Zone codes and land uses in Charleston county, 2016.

Class Code	General Definition	obs
101 - RESID-SFR	Single family residential	103,222
905 - VAC-RES-LOT	Vacant residential lot	24,644
160 - RESID-CNU	Residential condo (near Kiawah island)	15,581
120 - RESID-TWH	Residential (town houses)	8777
500 - General Commercial	General commercial	4,476
990 - UNDEVELOPABLE	Edges and awkward shapes	3,813
800 - AGRICULTURAL	Rural residential and resource management	3,131
952 - VAC-COMM-LOT	Vacant commercial zone	2,645
130 - RESID-DUP/TRI	Residential duplex/triplex	2,444
110 - RESID-MBH	Residential mobile home	1,722
750 - SPCLTY-REC	Agricultural residential/recreational	1,710
250 - SPCLTY-COMMCONDO	Residential condominium	1617
742 - HOA-PROP	Home owner association	889
691 - RELIGIOUS	Religious	684
210 - SPCLTY-SMA	Small apartment complexes	642
650 - SPCLTY-OFC	Office building	615
900 - RES-DEV-ACRS	Residential development acres	570
460 - AUTO-PARKING	Big lots with buildings that can be residential	563
220 - SPCLTY-TAMSBERG	Townhouses appraised as apartments	518
225 - SPCLTY-CNU-TMSBRG	Condos appraised as apartments	413
530 - SPCLTY-RTL	Retail	409
630 - SPCLTY-WHS	Warehouse	396
580 - SPCLTY-RST	Restaurant	355
200 - SPCLTY-APT	Large apartment complexes	323
671 - GOVT-BLDG	Government building	197
700 - SPCLTY-HTL	Hotel	178
140 - MH-PARKS	Mobile home park	171
195 - COMM-APP-RES	Residential use on commercially zoned land	160
681 - SCHOOLS	Schools	149
910 - COM-DEV-ACRS	Commercial development acres	142
170 - RESID-ROW	Residential rowhouse	97
481 - PUBLIC-UTIL	Public utility	91
304 - MFG/INDUST	Manufacturing/Industrial	64
121 - GROUP-LIV	Residential group living quarters	58
624 - CEMETERIES	Cemeteries	56
451 - ROAD-ROW	Roads	54
711 - MUSEUM-CULT	Public building	46
999 - Not Currently Classified	Not classified	40
300 - BUILDNG-ONLY	Rented buildings	38
471 - TELEPH-COMM	Company	16
411 - RAILRD/TRAIN	Train	9
150 - HOTELS	Hotel	3
10M-Mobile Home taxed as Real	Mobile home	1
Total		181,725

Table A.3: Zone codes and land uses in Berkeley county, 2016.

Zone	Type	Definition	obs
Berkeley County - Flex1	QR	Agricultural district	1,127
Berkeley County - R2	QR	Manufactured residential district	558
Moncks Corner - PD-R	QR	Residential Zone Districts	496
Summerville - R-4	QR	Multifamily residential district small-scale	388
Moncks Corner - R-1	QR	Single family residential district	313
Berkeley County - R2-R(F)	QR	Mobile home rural residential district	161
Moncks Corner - R-2	QR	Manufactured residential district	156
Moncks Corner - R-3	QR	Mobile home park	155
Berkeley County - R4	QR	Multifamily residential district small-scale	74
Berkeley County - R5	QR	Multifamily residential district (large-scale)	65
Berkeley County - R1-R	QR	Rural single-family residential district	60
Berkeley County - GC	QR	General Commercial district	37
Berkeley County - R1-MM	QR	Multisection manufactured residential district	23
St Stephen - TNR	QR	Temporary Neighborhood Residential	23
Moncks Corner - C-1	QR	Office Institutional District	15
Berkeley County - RNC	QR	Rural and neighborhood commercial district	12
Moncks Corner - TD	QR	Transitional District	11
Moncks Corner - C-2	QR	General Commercial District	10
Berkeley County - R15	QR	Preservation residential district	4
OUT	QR		4
Bonneau - AGRICULTURE	QR	Agricultural district	3
Berkeley County - R3	QR	Mobile home park	1
Bonneau - MOBILE HOME	QR	Mobile home park	1
Bonneau - SINGLE FAMILY	QR	Single family residential	1
St Stephen - HC	QR	Highway Commercial/Contiguous	1
Berkeley County - Flex1	AC	Agricultural district	28
Berkeley County - PD-MU	AC	Planned development /Mixed use	12
Summerville - I-1	AC	limited industrial	10
Berkeley County - HI	AC	Heavy industrial	1
Berkeley County - PD-OP/IP	AC	Office park/ Industrial park	1
Goose Creek - GC	AC	General Commercial district	1
Summerville - B-3	AC	General Business	1
Summerville - R-4	AC	Multifamily residential small-scale	1
Berkeley County - Flex1	AQ	Agricultural district	494
Berkeley County - R2	AQ	Manufactured residential	39
Berkeley County - HI	AQ	Heavy industrial	21
Berkeley County - R1	AQ	Single family residential	18
Moncks Corner - R-3	AQ	Mobile home park	13
Berkeley County - R15	AQ	Preservation residential	9
Berkeley County - GC	AQ	General Commercial district	8
Moncks Corner - PD-R	AQ	Residential Zone Districts	8
Berkeley County - R1-R	AQ	Rural single-family residential	5
Moncks Corner - C-2	AQ	General Commercial District	5
Goose Creek - PD	AQ	Planned District	4
Moncks Corner - R-2	AQ	Manufactured residential	4
St Stephen - TNR	AQ	Temporary neighborhood residential	4
Berkeley County - PD-MU	AQ	Planned development /Mixed use	3
Berkeley County - PD-OP/IP	AQ	Office park/ Industrial park	2
Berkeley County - R1-MM	AQ	Multi-section manufactured residential	2
Goose Creek - GC	AQ	General Commercial	2
Goose Creek - R-3	AQ	Mobile home park	2
Moncks Corner - C-1	AQ	Office institutional	2
Moncks Corner - PD-C	AQ	Planned district commercial	2
Berkeley County - R2-R(F)	AQ	Mobile home rural residential	1

Continued on next page

Table A.3 (continued).

Berkeley County - RNC	AQ	Rural and neighborhood commercial	1
Berkeley County - Flex1	OT	Agricultural district	1,870
OUT	OT		10
Bonneau - AGRICULTURE	OT	Agricultural	1
Berkeley County - R1	OT	Single family residential	1,052
Berkeley County - R2	OT	Manufactured residential	823
Berkeley County - PD-MU	OT	Planned development /Mixed use	600
Moncks Corner - R-3	OT	Mobile home park	355
Berkeley County - GC	OT	General commercial	296
Summerville - R-4	OT	Multifamily residential small-scale	210
Moncks Corner - R-1	OT	Single family residential	188
Berkeley County - R2-R(F)	OT	Mobile home rural residential	149
Moncks Corner - C-2	OT	General commercial	143
Moncks Corner - PD-R	OT	Residential	143
Moncks Corner - R-2	OT	Manufactured residential	116
Berkeley County - R5	OT	Multifamily residential (large-scale)	105
Berkeley County - RNC	OT	Rural and neighborhood commercial	53
St Stephen - TNR	OT	Temporary neighborhood residential	53
Berkeley County - R4	OT	Multifamily residential small-scale	50
Berkeley County - HI	OT	Heavy industrial	45
Summerville - B-3	OT	General business	45
Moncks Corner - C-1	OT	Office Institutional District	41
Berkeley County - R1-MM	OT	Multi-section manufactured residential	25
Berkeley County - R1-R	OT	Rural single-family residential	23
Berkeley County - PD-OP/IP	OT	Office park/ Industrial park	20
Berkeley County - LI	OT	Light industrial	19
Moncks Corner - TD	OT	Transitional	19
Berkeley County - R15	OT	Preservation residential	10
Summerville - I-1	OT	limited industrial	8
Summerville - PUD	OT	Planned unit development	8
Goose Creek - PD	OT	Planned	7
Berkeley County - OI	OT	Office institutional	6
Berkeley County - OIGC	OT	Office institutional /General commercial	6
Berkeley County - R3	OT	Mobile home park	6
Moncks Corner - M-1	OT	Light Industrial District	6
St Stephen - HC	OT	Light Industrial Office	6
Bonneau - MOBILE HOME	OT	Mobile home park	5
Goose Creek - R-3	OT	Mobile home park	4
Goose Creek - GC	OT	General Commercial district	3
Moncks Corner - MH-1	OT	Mobile home park	3
Moncks Corner - PD-C	OT	Planned District Commercial	3
Bonneau - SINGLE FAMILY	OT	Single family residential	2
St Stephen - LIO	OT	Light Industrial Office	2
St Stephen - M	OT	Light industrial	2
Summerville - R-6	OT	Multifamily residential	2
Bonneau - COMMERCIAL	OT	General commercial	1
St Stephen - TC	OT	Temporary commercial	1

Table A.4: Zone codes and land uses in Georgetown county, 2016.

Primary Land Use Code	Definition	obs
Q100	Improved residential lot	17,272
Q000	Vacant residential lot	7,240
Q650	Mobile home	6,528
Q200	Rural (less than 5 AC)	4,673
N450	Condominium	4,582
Q302	More than 5 AC with AG use	2,257
N65	Mobile home on property	2,046
N500	Improved resort lot	1,745
N300	Rural (more than 5 AC)	1,125
N455	Townhouses	774
N003	Vacant commercial	571
N451	Common area	488
E890	Church/Religious, etc	481
N700	Multiple lot discount value	475
N400	Vacant resort lot	316
N280	Office building	288
Q165	Mobile home	260
Q202	Rural (less than 5 AC with AG use)	245
E940	State	245
N456	Subdivision common area	240
Q657	MH on as storage	240
E930	County	239
N231	Retail Shop	233
N380	Marinas	214
E920	Veteran	207
E970	Georgetown water & sewer	200
E960 /Misc	187	
E452	HPR	173
N001	No infrastruature	168
N101	Improvement on lot	129
N050	Building only	120
Q009	Wetlands/ unbuildable	116
E950	City	114
N800	Duplex	110
N240	Warehouse	108
T600	Tax commission property	104
N150	Restaurant	72
N381	Boat storage	70
Q659	MH on as storage	68
N180	Convince store	67
N490	Garage/Storage	63
N370	Mobile home park	61
N651	Mobile home	55
E941	SC Public service	52
E910	Cemetery	44
N480	Storage warehouse	42
N701	Homeowners association	41
E980	Georgetown Hospital	40
N550	Medical Building	36
N004	Street, road or right of way	33
N453	Villa’s	33
N383	Boat slip	32
N260	Garages & Auto center	28
E931	Georgetown board of education	28

Continued on next page

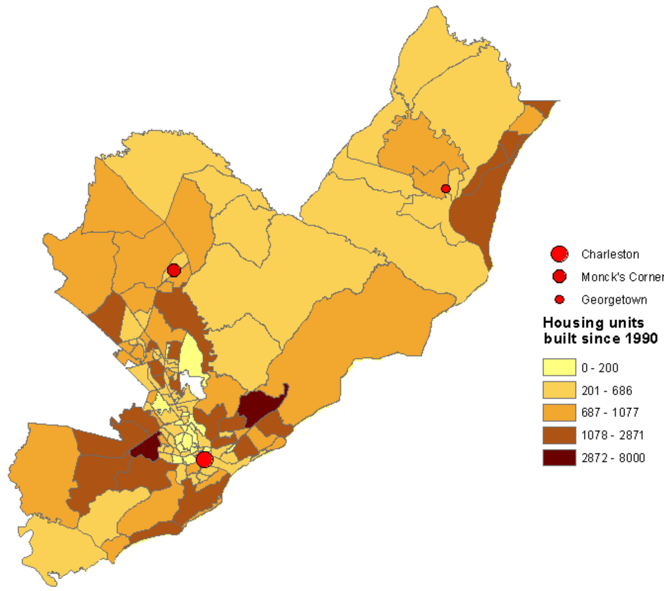
Table A.4 (continued).

N330	Banks	27
N440	Golf course	27
N460	Shopping center	27
E961	Federal government	26
N410	Apartments	25
N130	Club house	16
Q10	Unbuildable	15
N360	Barber/Beauty Shop	15
N250	Mini warehouse	14
N510	Auto center	14
N530	Auto dealer	14
Q653	Residential and commercial	14
N350	Service stations	13
N120	Motel	12
N190	Department store	12
N804	Quadplex	12
N340	Lumber Yard	11
N220	Discount store	10
N230	Shopping center	10
Q384	Boat at residence	10
N160	Fast food restaurant	7
N430	Day care center	7
E658	MH used with church	7
N170	Laundromat	6
N390	Nursery/Greenhouses	6
N570	Veterinary building	6
N140	Homes for elderly	5
N656	Commercial Property	5
Q652	Mobile Home on as R/E for Homestead	4
E990	PHONE, CABLE, POWER, ETC	4
N803	Triplex	3
N580	Fraternal building	2
N110	Hotel	1
N382	Boat ramp	1
Q385	Motor home/Ag 4%	1
N560	Convalescent hospital	1
Total		55,770

Table A.5: Zones of vacant lots in and out of the sample.

Land class	Vacant lots not in sample	Vacant lots in sample
Charleston county		
140 - MH-PARKS		151
110 - RESID-MBH		1,085
905 - VAC-RES-LOT	5,131	17,293
101 - RESID-SFR		84,013
800 - AGRICULTURAL	684	1,814
999 - Not Currently Classified	40	
910 - COM-DEV-ACRS	135	
952 - VAC-COMM-LOT	2,610	
990 - UNDEVELOPABLE	3,825	
Berkeley county		
Goose Creek - PD-MH (Planned Development Mobile Home Park)		2
Moncks Corner - MH-1 (Mobile home park)		5
Berkeley County - R3 (Mobile home park)		63
Berkeley County - R2-R(F) (Mobile home rural residential district)		661
Moncks Corner - R-3 (Mobile home park)		673
Goose Creek - R-3 (Mobile home park)		1,387
Public/Other (NA)	15,751	104
Georgetown county		
Mobile Home (Q650)		1
Mobile Home on as R/E for Homestead (Q652)		1
MH on as storage (Q659)		1
Lumber Yard (N340)		11
Mobile Home on as Real Estate (N651)		55
Mobile Home Park (N370)	29	59
No infrastructure (N001)	36	167
Rural (Less Than 5 AC with AG use) (Q202)	33	237
Mobile Home On as Real estate (Q165)	12	257
Vacant Commercial Property (N003)	30	556
Rural (More Than 5 AC) (N300)	2,043	998
More Than 5 AC with AG Use (Q302)	3,579	1,890
Mobile Home On Property (N065)	202	2,003
Rural (Less Than 5 AC) (Q200)	612	4,505
Vacant Residential Lot (Q000)	423	7,035

(a) New housing stock (1990–2016).



(b) Growth in housing stock (1990–2016).

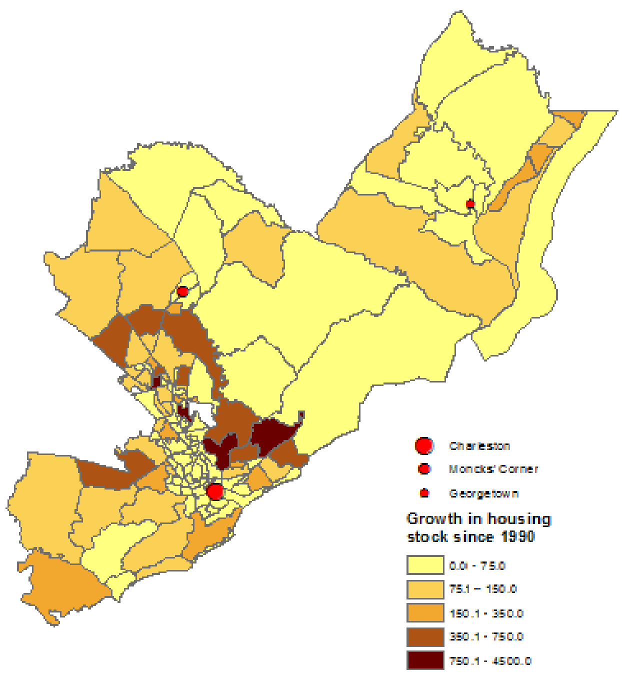


Figure A.1: Change in housing stock in the study area (1990–2016).

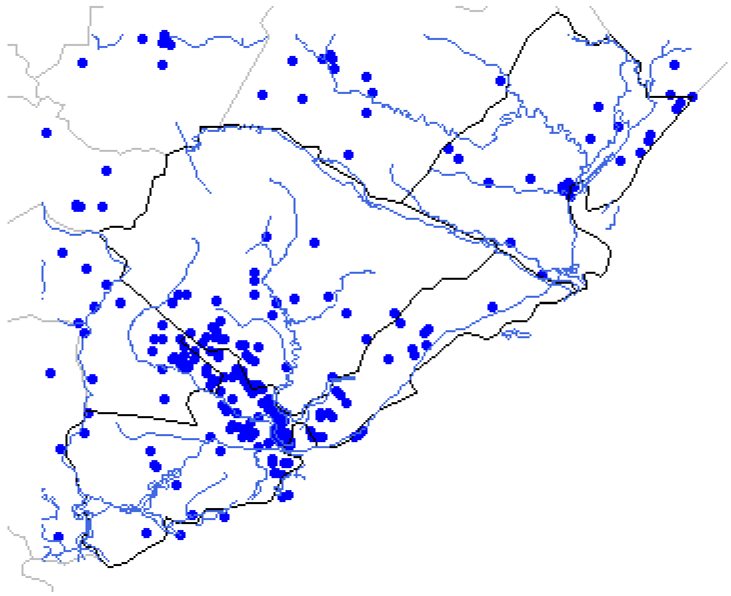
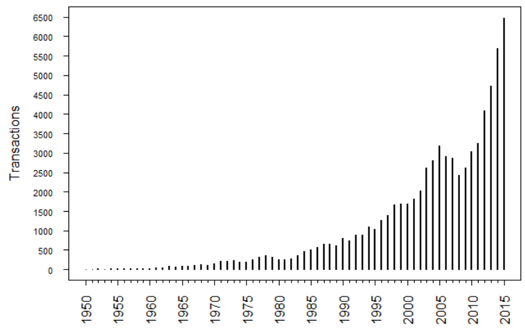
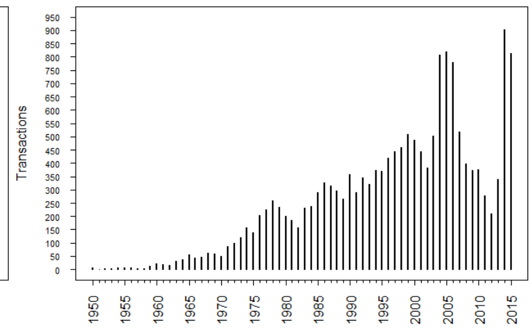


Figure A.2: Location of flash floods, floods, and coastal flooding incidents (1996–2016).

(a) Historical sales – Charleston.



(b) Historical sales – Berkeley.



(c) Historical sales – Georgetown.

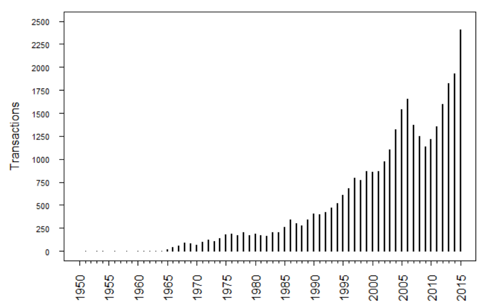


Figure A.3: Historical sales in study area (1950–2015).

B Modeling considerations: Are predicted values valid proxies?

A major interest of this study is to use the hedonics framework for predictive purposes and to generate valid proxies of land value that are then used as inputs in a second–stage probabilistic model of land use choice. Hence, the paragraphs below assess the accuracy of prediction attained in the first–stage estimation by analyzing the statistical properties of the generated proxies.[13] Tables B.1 and B.2 present different statistics to evaluate the predictive power of the models selected for estimation. Figures B.1 to B.4 illustrate in–sample and out–of–sample predictive power of those models.

To use valid proxies of land value in the model of land use change, it is important to specify hedonic models with high predictive power. With this focus, three type of modeling decisions were made: (1)

[13]Recall from section 2 that these proxies correspond to $\widehat{\pi^D}$ and $\widehat{\pi^U}$. Specifically, they are $\ln(\widehat{\frac{P^D}{Acre}})$, $\ln(\widetilde{\frac{P^D}{Acre}})$, $\ln(\widehat{\frac{P^U}{Acre}})$, and $\ln(\widetilde{\frac{P^U}{Acre}})$ in section 2.

decisions over which dependent variable to use and what functional form it should take (i.e., whether the hedonic models were to predict sales price or assessed values and whether these should be absolute or on a per acre basis, and whether linear, semi–log, or log–log functional forms should be used); (2) decisions about how to restrict the sample of parcels used in estimation (i.e., what time restrictions should be applied to the observed transactions data and how to define outliers in the sample); and (3) decisions over how to specify the models (i.e., whether to to use a set of variables from tax assessor data commonly available for all counties or a set of variables specific to each county).

To inform the making of these decisions, a standard machine learning algorithm known as the k–fold validation technique is employed.[14] In this application, k equals 10, and given tax assessor data availabilty, the modeling choices for all three counties were based on the results based on Charleston data. The output from the 10–fold cross validation is used to progressively refine baseline hedonic models for residential and agricultural parcels.[15] The baseline models are shown int he main text (equations 3.1 and 3.2).

In terms of the choosing a dependent variable, the model selection process indicates the preferred model is not better or worse at predicting sales prices than assessed values for either agricultural or residential parcels, thus sales prices are selected as the dependent variable, and to remain consistent with similar literature, price per acre was used as the dependent variable. Also, unsurprisingly, the logarithmic transformation of prices is more normally distributed that the non–transformed data. Thus, the functional form chosen for the hedonic models was log–linear. This choice also facilitates the interpretation of parameter estimates as semi–elasticities.[16]

[14]The k–fold cross validation technique is a technique used in statistics to asses how well the results of a model generalize to an independent data set. In k–fold cross-validation, the original sample is randomly partitioned into k equal size subsamples and the partitioning of data is done without replacement so that no observation in the data is found in two subsets. Of the k subsamples, a single subsample is kept as the validation data for testing the model, and the remaining $k-1$ subsamples are used as training data. (Training a model can be thought of as refining it. In machine learning, a model is trained using training data from a training set. To evaluate the trained model, the estimated coefficients are used to predict data in the test set. In my case, the test set is the same as the validation set.) The cross–validation process is then repeated k times (i.e., the number folds) with each of the k subsamples used exactly once as the validation data. The advantages of this method for model evaluation are well documented, the more salient being that all observations are used for both training (or refining) and validation, that each observation is used for validation exactly once, and that the k results from the folds can then be averaged (or otherwise combined) to produce a single estimation. An recent survey of cross-validation results is Arlot and Celisse (2010).

[15]The 10–fold cross validation technique results in 10 sets of predicted errors, each set is then manipulated to generate three measures of dispersion summarizing the quality of each estimation: the absolute deviations norm (commonly referred to as L1); the euclidean norm (or square root of the sum of squares norm, also known as $\sqrt{L}2$); and the the mean squared error (MSE). Each one of these three statistics offers different advantages when evaluating the quality of an estimator as they give proportionally higher or lower penalties to large errors or to predictions derived from small samples (for example, the $\sqrt{L}2$ measure will penalize large errors more heavily than the L1, and the MSE will penalize errors based on smaller samples more heavily than the $\sqrt{L}2$). To transparently select the statistical models that offer the better fit for the sample data, the distribution of L1, $\sqrt{L}2$, and MSE resulting from the validation process are examined.

[16]Within the literature, there is a history of swinging between advocating the use computationally feasible forms such as linear, semi–log, and log-linear, and favoring more general and flexible functional forms nested within the general quadratic Box–Cox form. The issue of complexity has been the focus of prior literature. For many years it seemed to be that simpler forms tend to perform best in the presence of omitted variables (Cassel and Mendelshn,

To inform the decision of restricting the use of historical data, the selection process was slightly different from the 10–fold cross validation approach previously described. Before of partitioning the data into 10 samples, the data was split into two subsamples, one with price data from properties that sold in 2015 and 2016, and one with price data from parcels that sold before 2015. The latter subsample (i.e., the pre–2015 data) was partitioned into 10 random samples, and each random sample was used as a training set to test the set with the 2015 and 2016 prices. Five historical subsets were tested: those with sales occurring between 1996 and 2014, 2000 and 2014, 2010 and 2014, sales after 1996 but excluding the years of the housing market crisis, and those after 2000 but excluding sales from 2006 to 2009. Results from the selection process suggest the preferred time frame for building predictive models of residential land value includes sales that occurred between 2010 and 2016. For the model of agricultural land value, it includes sales that occurred between 2000 to 2016 but exclude transactions during housing market crisis (defined from 2006 to 2009).

As with the decision over use of historical data, to inform the decision of removing outliers (i.e., unusually high or low prices), subsets of the data were created before partitioning the data into 10 random subsamples. Specifically, the data was split into five subsets containing different price ranges, each range incrementally more restricted than the previous. These ranges were defined to filter anomalous transactions that are known as "not at arm's length."In general, an arm's length transaction is one in which the buyers and sellers of a product act independently. For filtering purposes, out of arm's length transactions in the sample were defined as those for which the reported sales price differs greatly from the value estimated by tax assessors, and five restriction rules were made based on the ratio of these two values (both expressed in 2016 dollars).[17]

Five restriction rules are used to define unusable transactions. The internal validity of the model is assessed using each rule.[18] The selection mechanism indicates the preferred range of values for both models of land value (i.e., agricultural and residential), is the more restrictive of the five: the one with transactions were the ratio between prices and assessed values is between 0.5 and 1.5.

Finally, the process to guide the decision over which independent variables to include in the model results in similar but different specifications for each county. Results from the cross validation validation suggest that exploiting specificity of data on house characteristics at the county level does not necessarily offer an

1985; Cropper et al, 1988). However, in more recent years, other researchers have found that large gains in accuracy can be realized by moving from the standard linear specifications for the price function to a more flexible framework that uses a combination of spatial fixed effects, quasi-experimental identification, and temporal controls for housing market adjustment (Kuminoff, Parmeter, and Pope, 2010). Given the lack of consensus, the functional form is determined from the data (Palmquist, 2005).

[17]In the three counties studied, transactions that are not at arm's length are rather prevalent. Communication with local experts in tax assessor offices, GIS offices, and civic groups suggest that this anomaly is related to poorly defined property rights that are associated with local historical institutions such as the prevalence of "Heirs Property."

[18]The rules are: (1) no exclusion rule (i.e., all sales are included in the model); (2) exclude parcels whose corresponding sales price is less than 10% of the assessed value (or parcels with a price-to-assessed value ratio less than 0.1); (3) exclude parcels with a price–to–assessed value ratio less than 0.25; (4) exclude parcels with a price-to-assessed value ratio less than 0.5; and (5) exclude parcels with a price–to–assessed value ratio less than 0.5 and greater than 1.5 (or parcels whose price is less than 50% or over 150% of their assessed value).

advantage over using a general model with a set of variables common to all counties. Thus, a single baseline model is applied to all counties, and once a common set of property variables is chosen, further modifications are applied on the set of explanatory variables for each hedonic price function in each county.

In evaluating the internal validity of the models, it is found that, in general, the hedonic model for residential land value does not perform poorly at predicting in–sample values, and that it performs better than the hedonic model of agricultural land value. This is shown in table B.1.

Table B.1 presents different measures of prediction error for both models and shows how these statistics are substantially larger for agricultural models, with magnitudes that are between 4.5 to 12 times larger. Figure B.1 compares simple prediction errors (i.e., $Y - X\hat{\beta}$) for both models using boxplots and clearly shows that errors from the residential land model are more normally distributed with the mean near zero.

Figure B.2 compares the kernel densities of the original data to those of the predicted values and includes a measure of bandwidth for each kernel density which represents the standard deviation of the kernel. This comparison is shown for results of the residential model in panel (a) and of the agricultural model in panel (b). As shown in the figure, differences in bandwidths between kernels of original and predicted data are negligible for the residential model (0.007 instead of 0.09), while the difference between in bandwidths across agricultural values are rather pronounced (0.6 instead of 0.2). Panel (a) shows that, in general, the hedonic model for residential land value generates predictions that are closely related to in–sample values, while panel (b) shows that the agricultural model seems to largely misrepresent the group of parcels with lower prices—which is a particularly noise part of the sample (see reference to Heirs Property transactions). Nevertheless, predicted values for parcels with higher prices are do not differ drastically from the original data.

Figure B.3 shows the spatial distribution of predictive error. Darker colors indicate larger errors. Interestingly, the agricultural model tends to under–predict values disproportionately for parcels in Georgetown county and over–predict them elsewhere.

In terms of external validity, both hedonic models seem perform similarly for out–of–sample prediction. Table B.2 shows descriptive statistics for in–sample and out–of sample models of residential and agricultural parcels while figure B.4 juxtaposes kernel densities of in–sample and out–of–sample predicted values. As shown by table B.2 and figure B.4, the predicted out–of–sample values do not stand out as strikingly poor counterfactuals for either model, although again, there is some indication that residential and agricultural land models have slightly different predictive powers. Based on summary statistics and kernel distributions, in–sample predictions are slightly larger than out–of–sample predictions, however, that difference seems smaller for predictions of residential land value.

To summarize, there is indication that hedonic models for residential land have greater predictive power than models for agricultural land. This could be partly explained by the infrequency of transaction data for agricultural properties as well as by the greater uncertainty regarding their price is agreed on relative to residential parcels. However, both models generate out–of–sample predictions that are consistent with tendencies and distributions of in–sample estimates and original values.

B.1 Additional econometric considerations

B.1.1 Spatial dependence in the error term

The role of spatial processes is strong in determining how land is valued, how developers make decisions, and how natural landscapes evolve. When important covariates with spatial effects are not included in the model, or when the model is misspecified in another way, residuals will show spatial autocorrelation. For instance, consider the idea that proximity to sea walls determines the likelihood that a parcel gets flooded and therefore how landowners and developers value the parcel. If the location of sea walls is unknown, the effect of proximity to sea walls will be captured by the error term in a model of land value and the error term will exhibit spatial dependence. Specifically to this study, if an important covariate explaining spatial correlation among land values, development decisions, or the ecology of the landscape is not observed, spatial correlation in the error term can be a threat to the validity of any statistical inferences.

There are well supported ways to accommodate for spatial dependence in linear models, such as adding spatial weights to the error term (what is commonly known as using a spatial autoregressive error). However, the preferred solution for dealing with residual spatial autocorrelation is to find any missing covariates that explain spatial processes and include them in the model (Hoeting et al., 2006). Thus, before adding more structure to the error term in the hedonic models of land prices, and possibly introducing new randomness into the analysis, this study first adds spatial structure to the mean function by including multiple landscape attributes (for example the measures of surrounding land covers and measures of proximity) and spatial fixed effects at the census tract and block group level in the reduced form equation. Then, using standard diagnostic tools, the study tests whether adding spatial structure to the mean function is enough to account for the unobserved spatial autocorrelation in the process determining property values. Specifically, QQ–plots, bubble plots, spatial correlograms, and variograms are used to detect residual spatial autcorrelation in the hedonic models. Using these tools, it is found that including a rich set of spatial explanatory variables in the hedonic models is enough to remove any spatial dependence in the residuals.

In regards to modeling land use decisions, failing to address spatial correlation in land features that facilitate development could even result in biased coefficient estimates, and not just flawed standard errors (Bigelow, 2015). There are well documented ways to address the error dependence problem in a binary choice model.[19] However, there is evidence in the literature that in two–stage study, like the one conducted

[19]For instance, Lewis et al. (2011) address this issue by adding components to the random error term in the RUM

here, accounting for spatial correlation in the first–stage can be enough to fully capture spatial processes that influence the second–stage process, therefore obviating the need to use complicated techniques that make a model more complex and the analysis more prone to error (Newburn et al., 2006).

Furthermore, to the extent that land use patterns and the spatial configuration of aggregated development decisions are perceived as disorganized and unpredictable, and that the study of individual landowner conversion decisions appears insufficient for understanding sprawl, there may not be any gain from separately addressing spatial autocorrelation in the second stage of the analysis.[20] Therefore, for the purpose of this study, relying on results of the diagnostics tests for the hedonic estimations and including in the land price equations important spatial features thought to influence the formation of aggregated pockets of development, such as proximity to large urban centers and presence of development in neighboring areas, is considered enough to address potential spatial autocorrelation in the development decision. Statistical strategies to model spatial processes are presented in appendix D. That discussion also includes a presentation of the tools used to diagnose spatial autocorrelation in my application and the results that support the conclusion of continuing without further corrections.

B.1.2 Sample selection correction

By necessity, the hedonic model of residential property value is estimated using observed sales prices of parcels that have been converted to residential use. Thus, there is an inherent selection bias that threatens the estimation of coefficients. To estimate valid (i.e., unbiased) coefficients in a hedonic model for developed parcels, the corrective procedure proposed by Heckman (1979) can be used. This procedure calls for the identification of at least one exclusion variable (i.e., a variable that affects the probability that parcels develop but not the sales price of residential parcels).

To implement Heckman's sample selection corrective method to my empirical application, the exclusion restriction variables are measures soil quality and proximity to agricultural processing facilities, both factors that are expected to affect agricultural land value but not residential land value. Despite the theoretical

that allow land values to be correlated spatially (i.e., all parcels within a county share a common value for a given term) and temporally (i.e., each parcel has a common value for another term across time periods). An alternative way to address spatial dependence is to extract subsamples of the data that are less likely to be spatially correlated with each other, and then estimate the land use model on these subsamples. The justification behind this approach is that the spatial autocorrelation in the residuals is likely to be lower if the samples used for estimation are farther apart in a spatial sense. There are multiple options for selecting the samples that are to be included in the estimation. For example, Carrion-Flores and Irwin (2004) follow the so-called "work-around"method, creating a subsample of the data by removing nearest neighbors within a fixed distance. In a different study, Newburn et al. (2006) perform multinomial logit regression on random stratified bootstrapped samples taken from the full dataset to correct the error term. More recently, Bigelow (2015) constructs a block-stratified random sample and uses it to estimate the econometric land use models.

[20]Disproportionate growth of urban areas and excessive leapfrog development are examples of such inefficient patterns, and the term to describe this inefficiency is "sprawl."Several hypotheses to explain sprawl have been tested but empirical evidence of what factors influence sprawl per se remains limited (Carrion–Flores and Irwin, 2004).

strength behind the logic used to select these instruments, a close examination of the data offers weak support for the choice of instruments, questioning whether any instrument should be used at all. Evidence shows that using weak instruments is likely to produce inconsistent and sometimes biased (Bound, Jaeger and Baker, 1995). Thus, to minimize introduced damage in the estimation and avoid the risk of over-fitting the model, this study does not pursue any corrective steps in the final estimation of the hedonic model for residential properties.

Other works in the literature have opted to proceed with the non–corrective method (Wrenn and Irwin, 2012; Newburn et al., 2004-2006; and Bockstael, 1996). In fact, few works empirically address the issue of sample selection bias, even though the literature clearly recognizes the problem. A theoretical justification for abstracting away from the sample selection problem is the idea that a hedonic model of residential land value may be unbiased if developers' ignorance of omitted variables that drive the selection problem matches the ignorance of the researcher (Bockstael, 1996). Appendix E includes findings from the exploratory analysis of instruments for a selection bias correction, and to illustrate the trivial effect of including or excluding these weak instruments from estimation, the results from hedonic regressions that correct for sample selection bias using the proposed weak instruments and the Heckman methodology are presented.

Table B.1: Measures of prediction error from in–sample estimation of hedonic models.

	Agricultural parcels	Residential parcels
$L1 = \sum \lvert Y - \hat{Y} \rvert$	214,996.1	17,708.5
$\sqrt{L2} = \sqrt{\sum (Y - \hat{Y})^2}$	1,452.92	379.72
$MSE = \frac{1}{N} \sum (Y - \hat{Y})^2$	63.13	13.93

Table B.2: Summary statistics: in–sample against out–of–sample.

	Residential		Agricultural	
	in–sample	out–of–sample	in–sample	out–of–sample
	$\ln(\widehat{\frac{P^D}{Acre}})$	$\ln(\widetilde{\frac{P^D}{Acre}})$	$\ln(\widehat{\frac{P^U}{Acre}})$	$\ln(\widetilde{\frac{P^U}{Acre}})$
min	9.85	8.25	6.69	7.28
1st Quantile	13.49	12.67	10.55	10.55
Median	13.88	13.51	11.81	11.2
Mean	13.91	13.38	12.57	11.87
3rd Quantile	14.25	14.18	14.56	12.13
Max	16.81	16.82	19.48	19.38

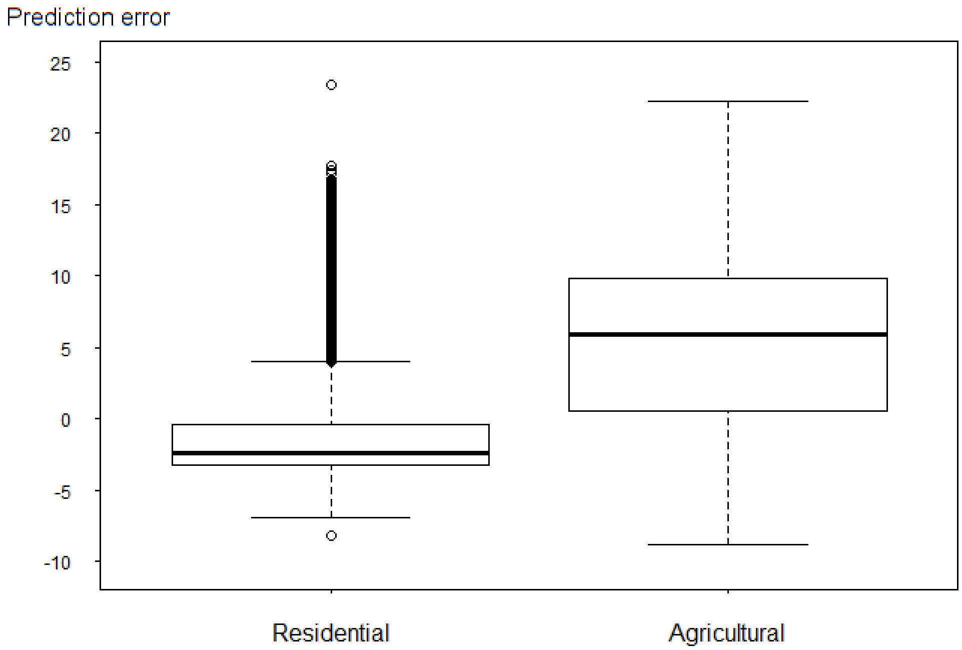
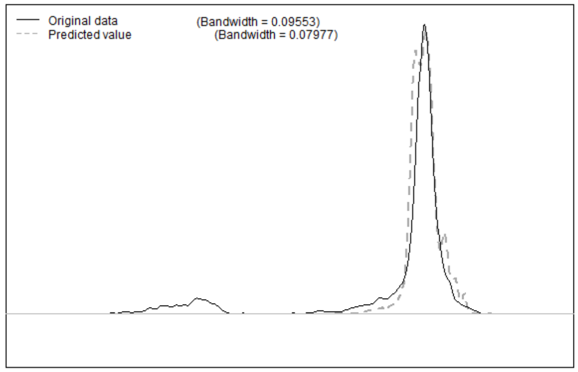


Figure B.1: Boxplots of in–sample prediction error of parcels in the sample. In the boxplot, the thick middle line corresponds to the median, and the edges of the box to the first and third quartiles. For the residential model, whiskers are the inner fence (1.5 times the interquartile range above the upper quartile and below the lower quartile), and the dots are outliers. For the agricultural model, whiskers and the minimum and maximum observations.

(a) Predictions from the hedonic model of residential land value.



(b) Predictions from the hedonic model of agricultural land value.

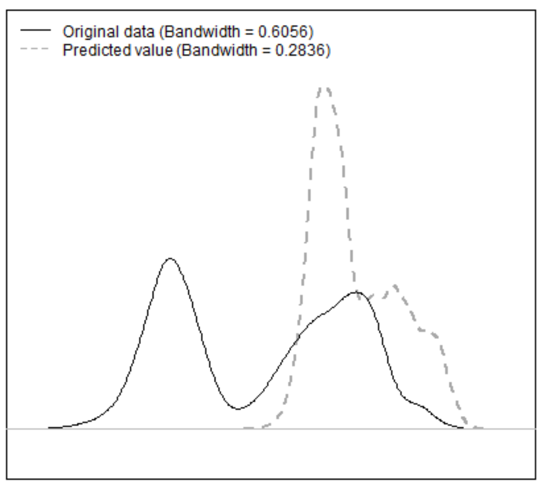
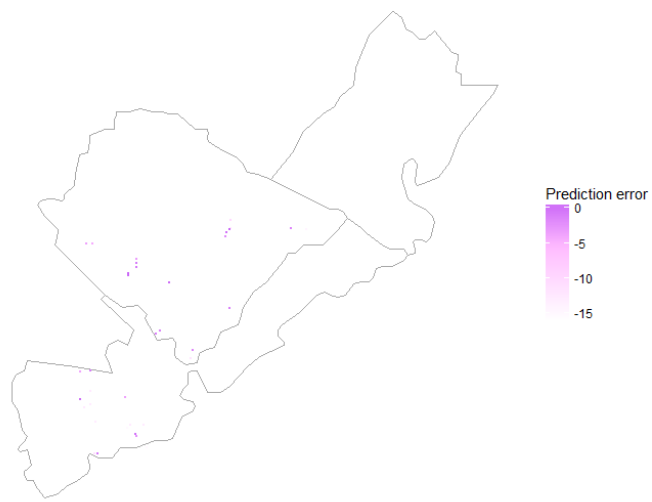


Figure B.2: Densities of observed and predicted land values for parcels in the sample.

(a) Predictions from the hedonic model of residential land value.



(b) Predictions from the hedonic model of agricultural land value.

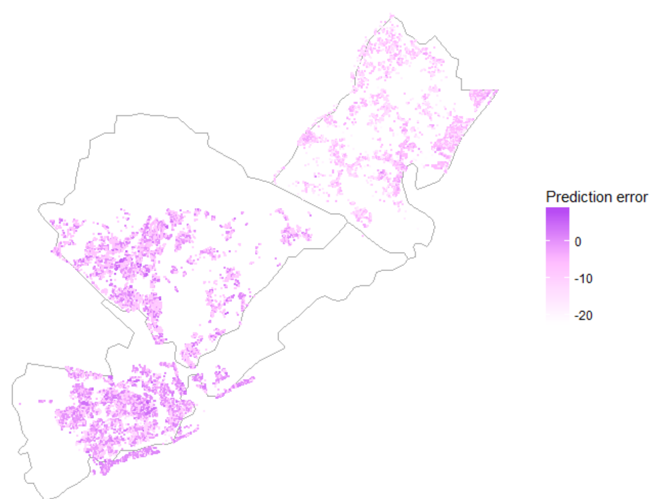
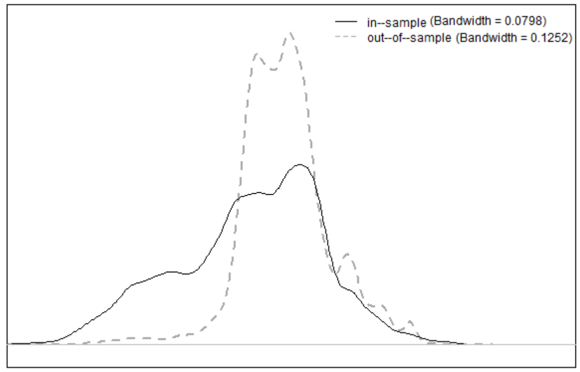


Figure B.3: Prediction error of land values for parcels in sample.

(a) Predictions from the hedonic model of residential land value.



(b) Predictions from the hedonic model of agricultural land value.

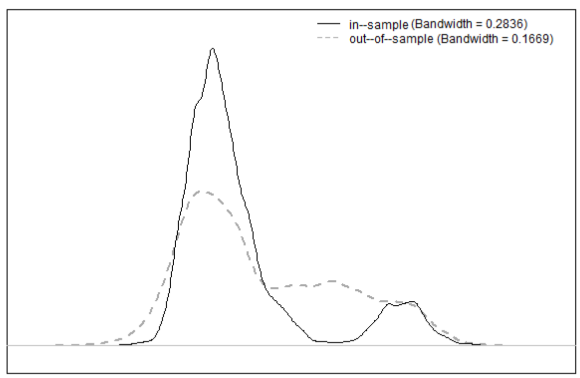


Figure B.4: Densities of in–sample and out–of sample prediction of land value.

C Output from 10–fold validation

In the empirical analysis, a repeated cross–section database is build using the most recent recorded sales price for each property in the sample (i.e., last–price data), and to select the model that best fits this data, this study follows a 10–fold cross validation technique and examine different measures of error (L1, L2, and MSE) to determine which regressors to include in the model and to guide the removal of noisy observations. Below, output figures used to arrive at the final choice of model are included.

C.0.0.1 Output from 10–fold cross validation procedure used for guiding the selection of the hedonic model for agricultural land value

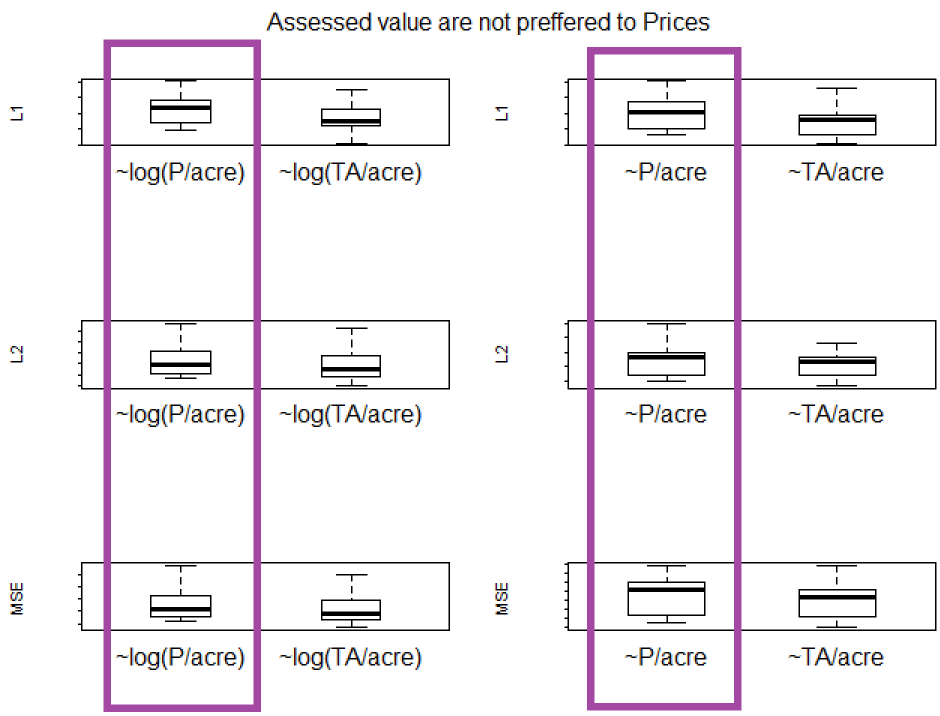


Figure C.1: Dependent variable choice (price vs. assessed values).

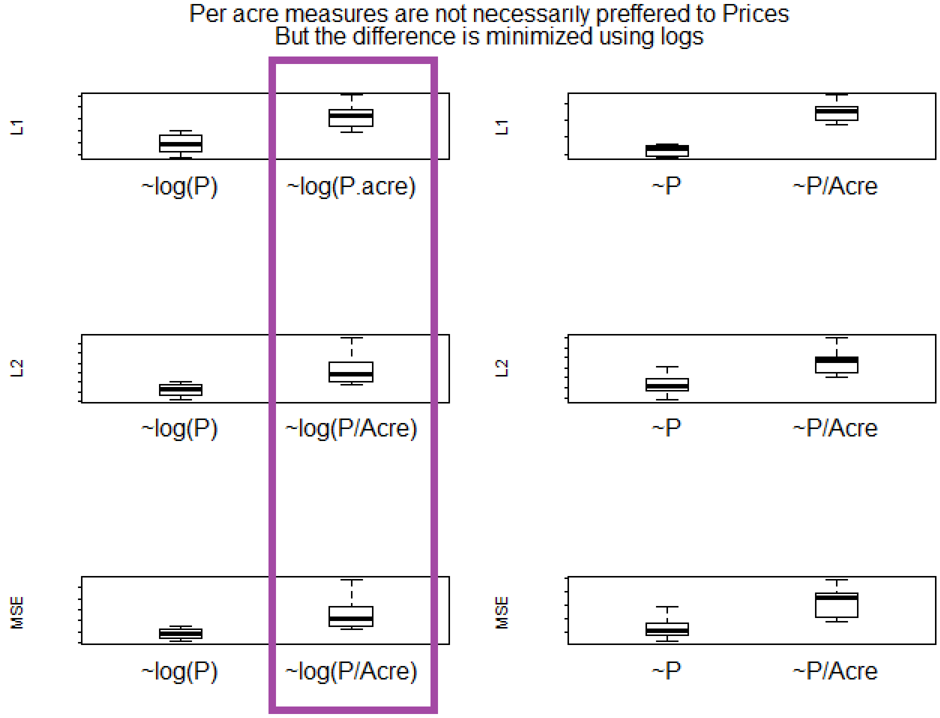


Figure C.2: Choice of dependent variable form (log vs levels).

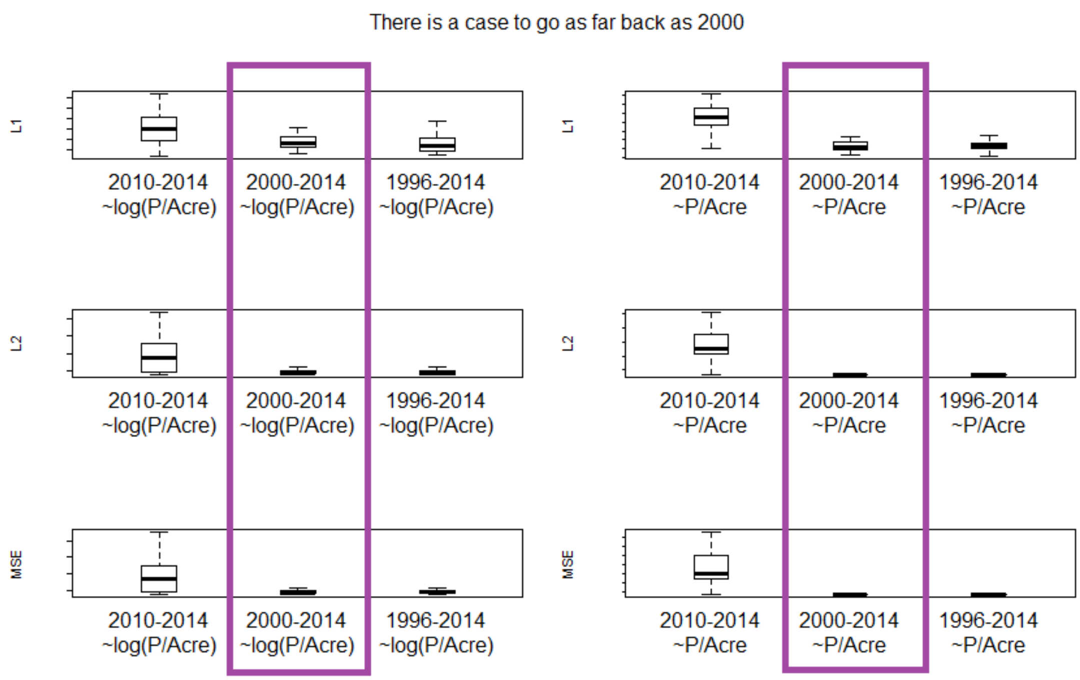


Figure C.3: Choice of historical data.

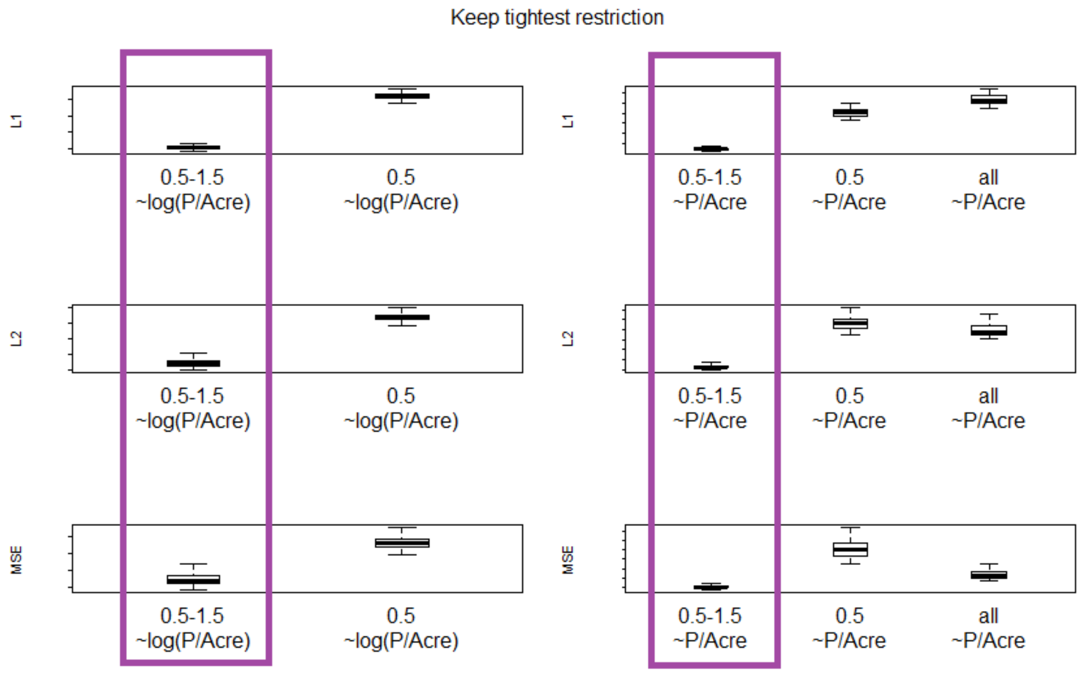


Figure C.4: Price range restriction.

C.0.0.2 Output from 10–fold cross validation procedure used for guiding the selection of the hedonic model for residential land value

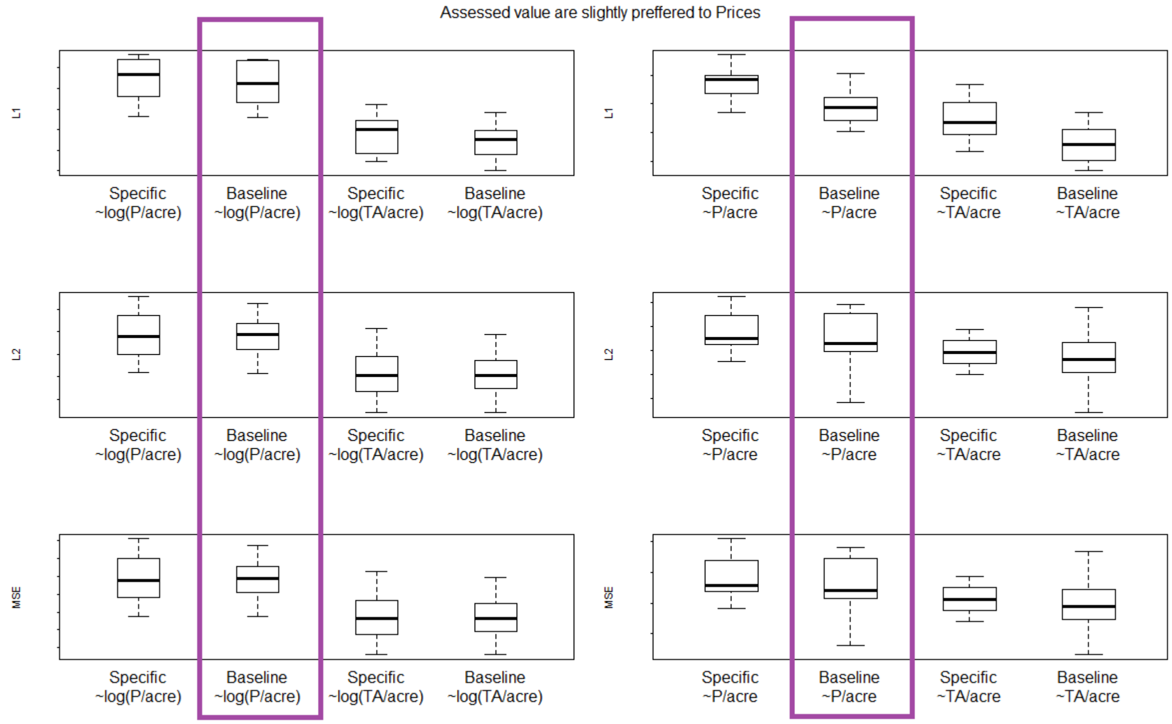


Figure C.5: Dependent variable choice (price vs. assessed values, logs vs. levels).

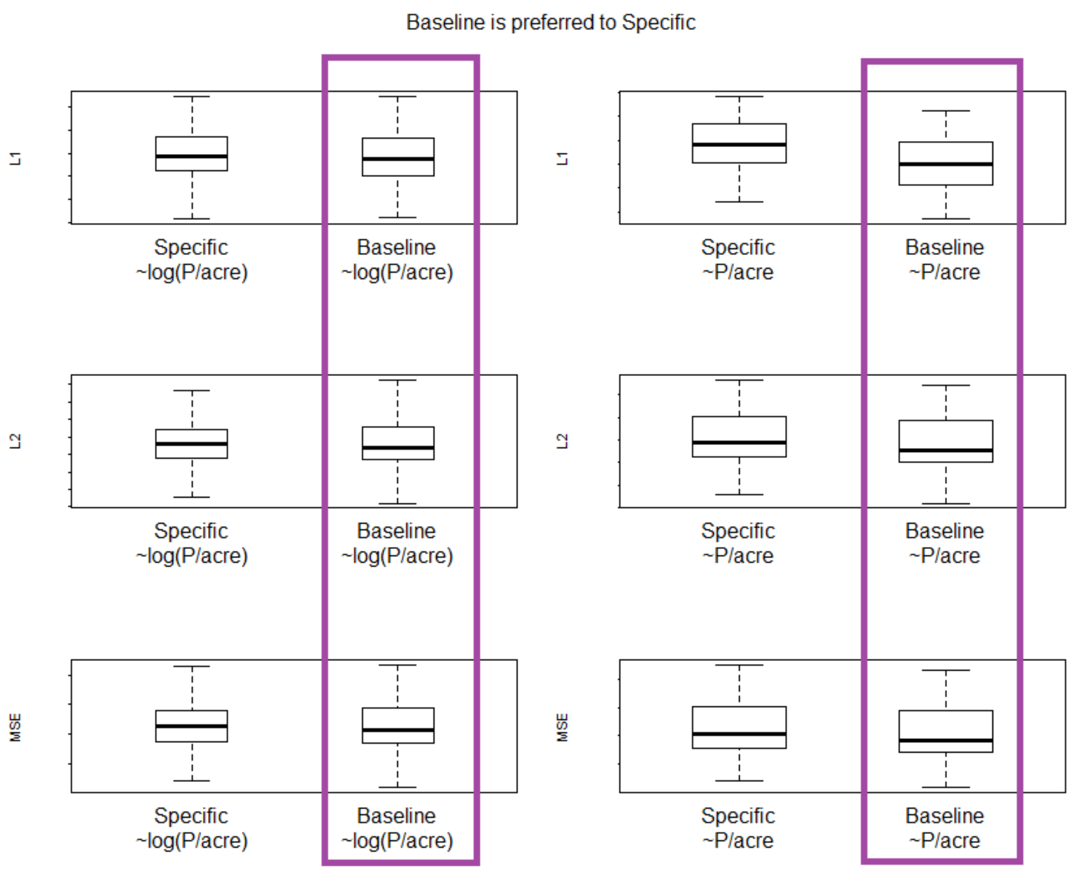


Figure C.6: Choice of county specific variables.

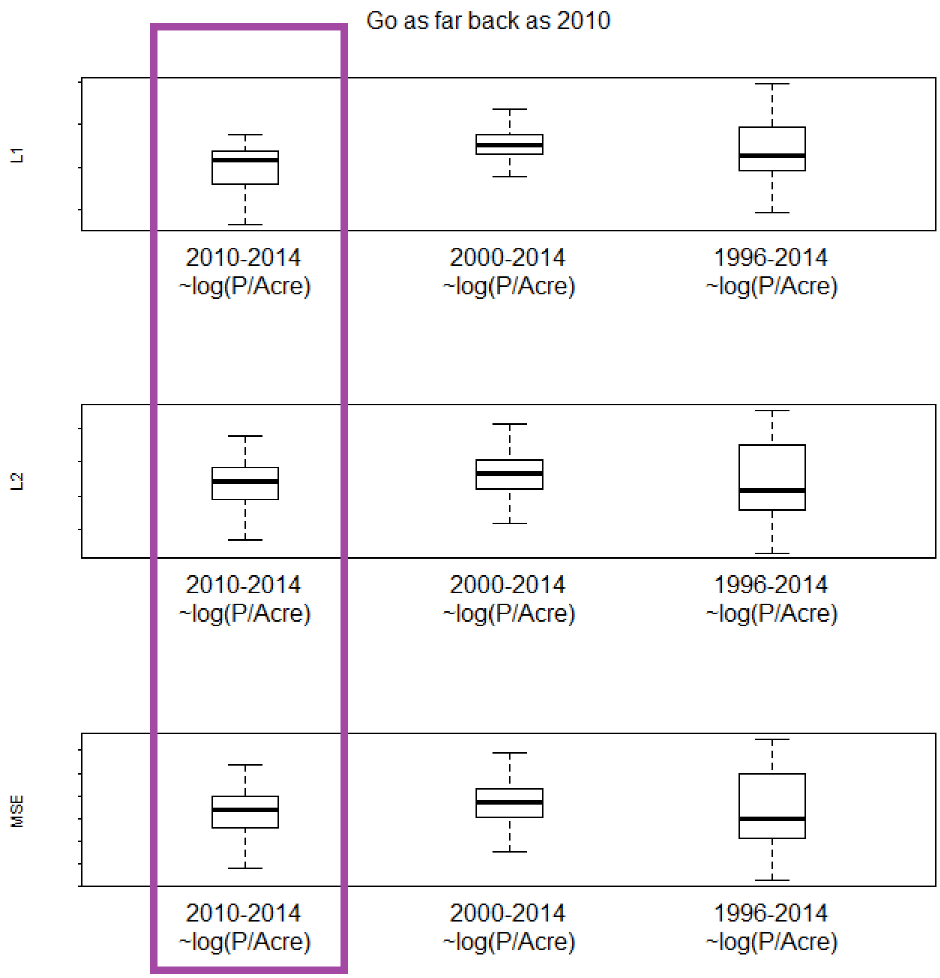


Figure C.7: Choice of historical data.

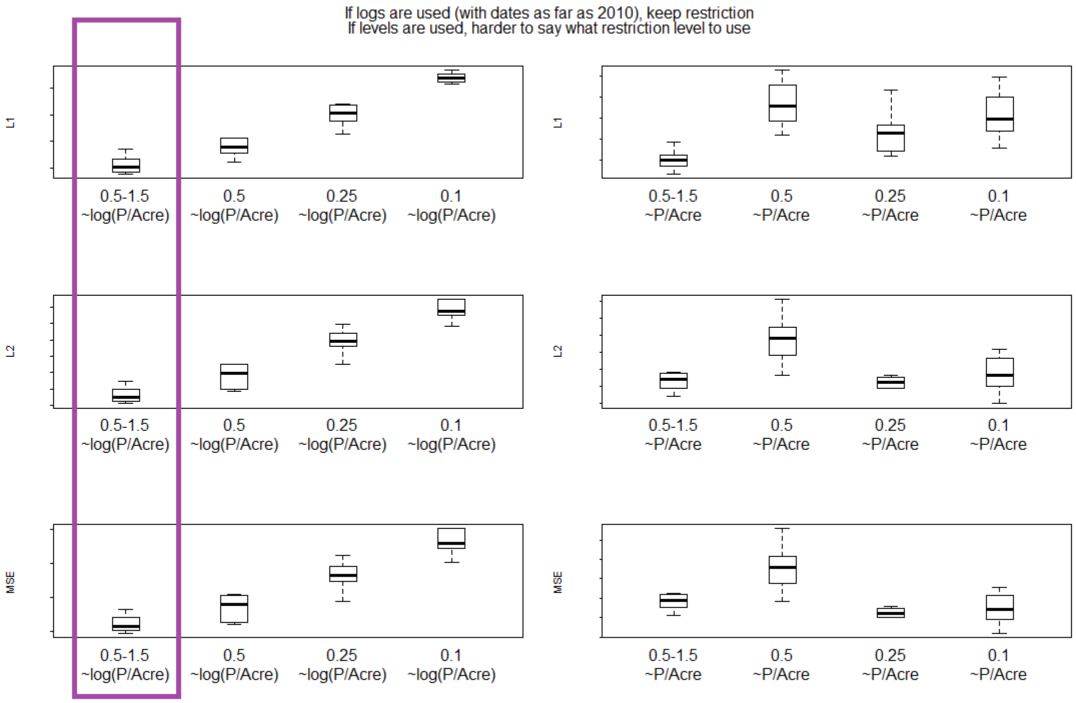


Figure C.8: Price range restriction.

D Spatial dependence diagnostics

The role of spatial processes is strong in determining how land is valued, how developers make decisions, and how natural landscapes evolve. Land use decisions and land values may be correlated across space if there are gradual changes in local or regional features, such as distance to processing facilities and soil quality. In turn, when it comes to environmental processes, the ecological value of a given parcel may be contingent on the land use status of neighboring parcels. For instance, when concerned with managing wetlands for flood protection, the ecological value of a given parcel may not justify the costs of protection if the parcel is surrounded by parcels that have already been developed.

When important covariates with spatial effects are not included in the model, or when the model is misspecified in another way, residuals will show spatial autocorrelation. For instance, consider the idea that proximity to sea walls determines the likelihood that a parcel gets flooded. If the location of sea walls is unknown, the effect of proximity to sea walls will be captured by the error term in a model of land value and the error term will exhibit spatial dependence. Spatial autocorrelation in land features that facilitate development can be a threat to the validity of any statistical inferences. Thus, spatial relationships must be accounted for in the econometric analysis.

When the hedonic models are specified as spatial processes, the price models are:

$$\begin{aligned} P_{it}(s) &= \alpha_i + X_{it}\beta + \nu_{it} \\ &= \alpha_i + \mu_{it}(s) + \eta_{it}(s) + \epsilon_{it}, \end{aligned} \tag{D.1}$$

where $X_{it}\beta$ constitute what in spatial statistics is known as the mean function $\mu_{it}(s)$, which explains large–scale spatial variation, while small–scale spatial variation is captured by the error term, ν_{it}. The latter can be decomposed as $\nu_{it}(s) = \eta_{it}(s) + \epsilon_{it}$. Where $\eta_{it}(s)$ captures spatial variation in the residuals, and ϵ_{it} is the uncorrelated error term.

If $\eta_{it}(s)$ is unobserved, then, its effect on prices is captured by $\nu_{it}(s)$, and ν_{it} exhibits spatial dependence. Also,$\nu_{it}(s)$ would be correlated with unobservables in $\eta_{it}(s)$.

In linear models, like the hedonic models in this study, uncorrected spatial dependence in the error term does not necessarily lead to bias of the coefficients but affects the standard errors. Thus, failing to account for local spatial processes driving property prices may result in inaccurate inference.

In non–linear models, like binary choice models of land use, spatial correlation may even result in biased coefficient estimates (Bigelow, 2015).[21] However, in a two–stage process like the one conducted here, it is

[21] There are well documented ways to address the error dependence problem in a choice model. For instance, Lewis et al. (2011) address this issue by adding components to the random error term in the RUM that allow land values to be correlated spatially (i.e., all parcels within a county share a common value for a given term) and temporally

possible that correcting for spatial correlation in the first–stage models is enough to fully capture spatial processes in the second–stage process, therefore obviating the need to use complicated techniques that make a model more complex and the analysis more prone to error (Newburn et al., 2006).

There are well supported ways to accommodate for spatial dependence for linear models, such as adding spatial weights to the error term (what is commonly known as using a spatial autoregressive error). However, the preferred solution for dealing with residual spatial autocorrelation is to find any missing covariates that explain spatial processes and include them in the model (Hoeting et al., 2006). Thus, before adding more structure to the error term in the hedonic models of land prices and possibly introducing new randomness into the analysis, this study first adds spatial structure to the mean function by including multiple landscape attributes (for example the measures of surrounding land covers) and spatial fixed effects at the census tract and block group level in the reduced form equation. Then, using standard diagnostic tools, the study tests whether adding spatial structure to the mean function is enough to account for the unobserved spatial autocorrelation in the process determining property values. Specifically, QQ–plots, bubble plots, spatial correlograms and variograms to detect residual spatial autcorrelation in the hedonic models are used. Using these tools, this study finds that including a rich set of spatial explanatory variables in the hedonic models is enough to remove any spatial dependence in the residuals.

D.1 Visual inspection

To study whether or the estimated hedonic models are appropriate for representing the spatial process determining land values in the sample, this study first makes a visual exploration of the residuals. Standard QQ–plots showed that residuals from models rich in spatial variables were in general normally distributed, suggesting the models are appropriate for the data in the sample and that it is not inadequate to proceed to make inference from these models. This conclusion was supported by results from complementary "bubble plots."[22] These plots showed that positive and negative residuals, large or small, do not seem to be clustered nor perfectly dispersed, suggesting that the independence assumption made in the standard regression

(i.e., each parcel has a common value for another term across time periods). An alternative way to address spatial dependence is to extract subsamples of the data that are less likely to be spatially correlated with each other, and then estimate the land use model on these subsamples. The justification behind this approach is that the spatial autocorrelation in the residuals is likely to be lower if the samples used for estimation are farther apart in a spatial sense. There are multiple options for selecting the samples that are to be included in the estimation. For example, Carrion-Flores and Irwin (2004) follow the so-called "work-around"method, creating a subsample of the data by removing nearest neighbors within a fixed distance. In a different study, Newburn et al. (2006) perform multinomial logit regression on random stratified bootstrapped samples taken from the full dataset to correct the error term. More recently, Bigelow (2015) constructs a block-stratified random sample and uses it to estimate the econometric land use models.

[22]Bubble plots are an intuitive informal diagnostic tool provided by the package gstat in the statistical software R. They plot the residuals against their spatial coordinates. The size of the dots is proportional to the value of the residuals, and different colors are used for positive and negative residuals.

model is met (i.e that residuals are randomly distributed). This was observed for models on residential and agricultural properties in all three counties.

D.2 Spatial correlograms

A more formal way of detecting residual spatial autocorrelation is a spatial correlogram. A spatial correlogram is a plot of the Moran's Index, or Moran's I, as a function of distance. Moran's I is computed with following formula:

$$I = \frac{n}{W} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n} (y_i - \bar{y})^2},$$

where n equals the total number of parcels in the study area, y_i is the land value of parcel i, $\bar{y}$ is the mean land value in the sample, w_{ij} represents the spatial weights, and W is the sum of all weights in the study area. When using Moran's I, the null hypothesis is that there is no spatial correlation and it can be tested using a permutation test. [23]

The way the weights w_{ij} are defined has a strong influence on the actual value of I. In the simplest case, the weights have the values 1 when parcels i and j are neighbors or 0 when they are not neighbors. However, there are more possible definitions what makes two parcels neighbors. For instance, one option is to consider two parcels as neighbors ($w_{ij} = 1$) if they are within 0.5 miles or 0.25 miles from each other. A spatial correlogram plots values of I as a function of the distance used to define the weights, w_{ij}'s.[24]

What is found from using spatial correlograms, is that there is no strong autocorrelation in the residuals for any of the distances considered for defining the weights.[25] Furthermore, starting at a distance of 0.3 miles, land values of neighboring parcels clearly appear to be independent from each other.

[23]If there is no spatial autocorrelation, the expected value of I is $E[I] = -1/(n-1)$, which is close to 0 when n tends to infinity. When $E[I] = -1$, the process exhibits perfect dispersion, that is, similar values repel each other. Instead, when $E[I] = +1$, the process exhibits perfect clustering. A permutation test is a statistical tool similar to bootstrapping where the observed data is resampled a large number of times to generate a distribution from which to assess whether or not the tested effect is random. Unlike boostrapping the resampling in a permutation test is done without replacement.

[24]To produce a meaningful correlogram, the width of the distance bands should be greater than the minimum distance between parcels in the study area. The correlog function of the ncf package in R allows computing a spatial correlogram and tests the significance of the different I values with permutation. Results from hypothesis testing using the correlog function include plots indicating when values of I significantly larger or smaller than expected under the null hypothesis of no autocorrelation. Black dots in the plot indicate that the corresponding values of I are significantly different from $-1/(n-1)$.

[25]Morgan's I's corresponding to weights set on distances between 0.3 and 60 miles with 0.3 mile increments are tested. The spatial correlograms show intermitently significant although weak autocorrelation in the residuals (i.e., I's are close to zero). But because the values are so close to zero, it is difficult to conclude there is a case of autocorrelation.

D.3 Spatial variograms

Another common tool used in geostatistical data analysis used to detect residual spatial autocorrelation are variograms (or semi-variogram). Variograms are plots of the semivariance as a function of distance. The semivariance of a function expresses the degree of relationship between parcels on a surface and it equals half the variance of the differences between all possible parcels spaced at a constant distance apart. The semi-variance between two sites is defined by:

$$\gamma(x_1, x_2) = \frac{1}{2}E[(Z(x_1) - Z(x_2))^2],$$

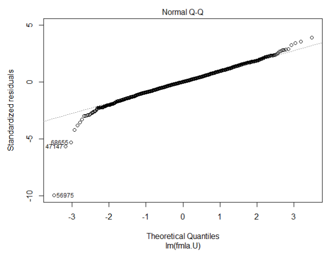
where x_1 and x_2 are the coordinates of the parcel and Z is a random function modeling the process of how land values are determined. If there is strong spatial dependence, parcels that are closer together show a smaller semivariance.

What is found from examining variograms is that the semi-variance does not seem to change with distance, suggesting that parcels are not more or less similar to parcels that are close to each other than they are to parcels that are more distant, therefore supporting the assumption of no spatial autocorrelation.[26]

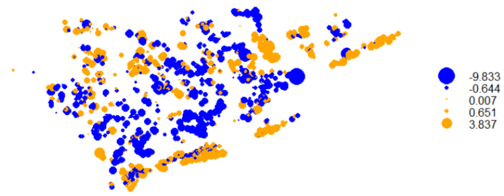
Having performed these three types of tests, this study concludes that that spatial autocorrelation is not present in the residuals from the selected hedonics models and therefore no further steps are taken in the structural correction of residuals in the empirical analysis.

[26]The gstat package in R also has functions for variogram analyses. Namely, two types of variogram tools are used: variogram clouds and experimental variograms. A variogram cloud plot the semi-variance as a function of distance for every possible pair of parcels. These plots can be difficult to interpret due to the massive density of possible parcel pairs. Thus, to improve the ease of analysis, experimental variograms, which are like variograms that group distances into distance classes, are also used.

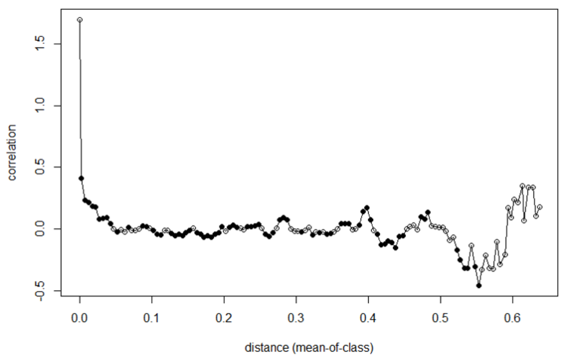
(a) Standard QQ plot showing that residuals are close to normally distributed.



(b) Bubble plot. The size of the dots is proportional to the value of the residuals, and different colors are used for positive and negative residuals.



(c) Spatial correlogram. It plots Moran's Index as a function of distance (1 in x-axis = 60 miles).



(d) Experimental variogram. It groups distances into classes and displays the average semi-variance among each class (1 in x-axis = 60 miles).

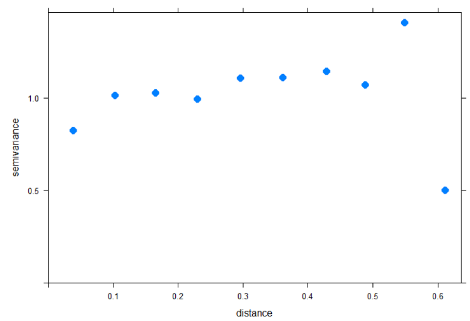
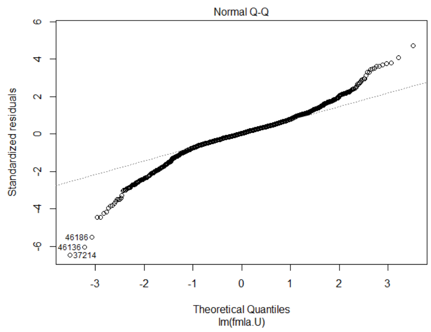
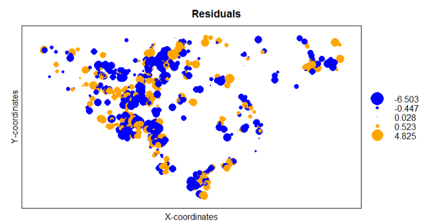


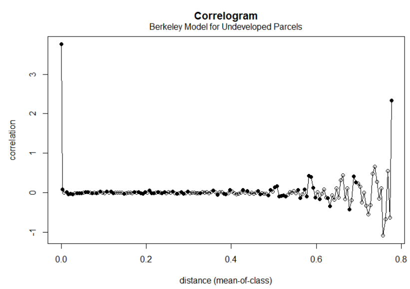
Figure D.1: Diagnostic tools used to visually assess existence of spatial dependence in the error term of the hedonic model of undeveloped parcels in Charleston county.



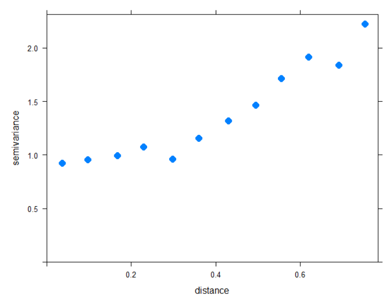
(a) Standard QQ plot showing that residuals are close to normally distributed.



(b) Bubble plot. The size of the dots is proportional to the value of the residuals, and different colors are used for positive and negative residuals.



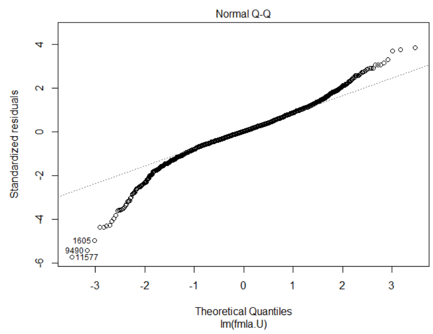
(c) Spatial correlogram. It plots Moran's Index as a function of distance (1 in x-axis = 60 miles).



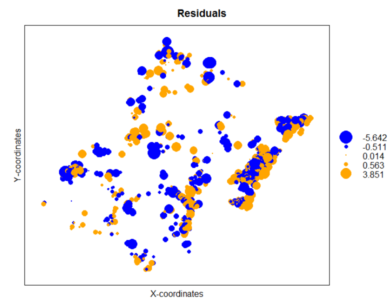
(d) Experimental variogram. It groups distances into classes and displays the average semivariance among each class (1 in x-axis = 60 miles).

Figure D.2: Diagnostic tools used to visually assess existence of spatial dependence in the error term of the hedonic model of agricultural parcels in Berkeley county.

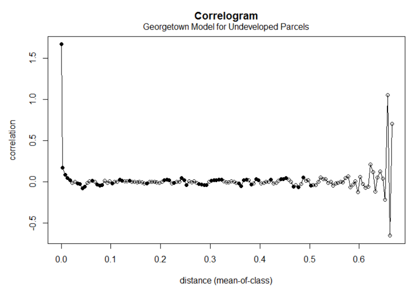
(a) Standard QQ plot showing that residuals are close to normally distributed.



(b) Bubble plot. The size of the dots is proportional to the value of the residuals, and different colors are used for positive and negative residuals.



(c) Spatial correlogram. It plots Moran's Index as a function of distance (1 in x-axis = 60 miles).



(d) Experimental variogram.It groups distances into classes and displays the average semi-variance among each class (1 in x-axis = 60 miles).

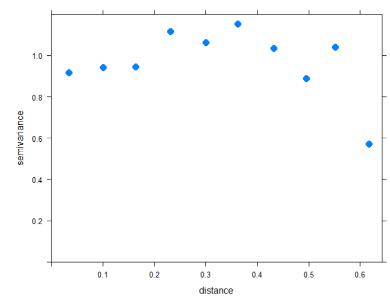
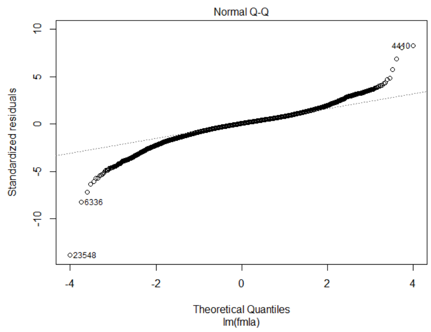
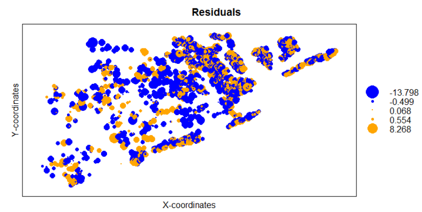


Figure D.3: Diagnostic tools used to visually assess existence of spatial dependence in the error term of the hedonic model of agricultural parcels in Georgetown county.

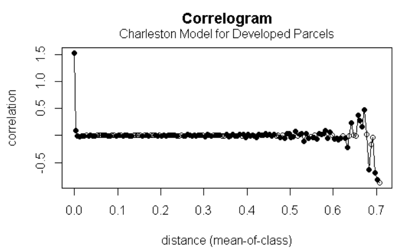
(a) Standard QQ plot showing that residuals are close to normally distributed.



(b) Bubble plot. The size of the dots is proportional to the value of the residuals, and different colors are used for positive and negative residuals.



(c) Spatial correlogram. It plots Moran's Index as a function of distance (1 in x-axis = 60 miles).



(d) Experimental variogram.It groups distances into classes and displays the average semi-variance among each class (1 in x-axis = 60 miles).

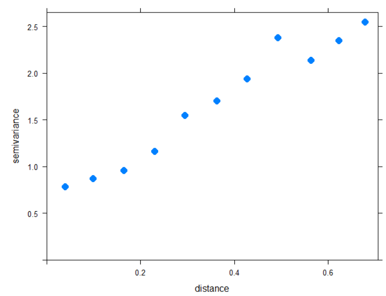
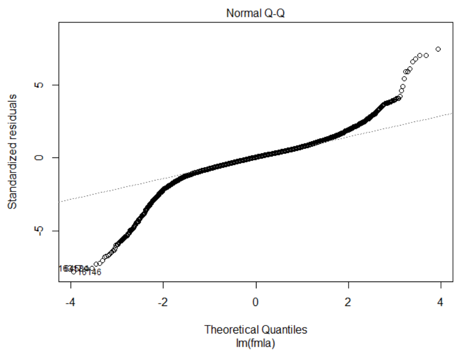
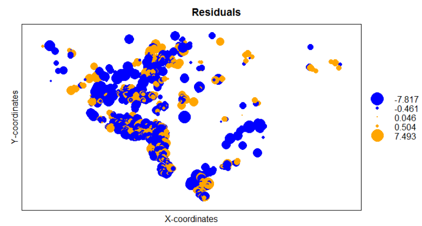


Figure D.4: Diagnostic tools used to visually assess existence of spatial dependence in the error term of the hedonic model of residential parcels in Charleston county.

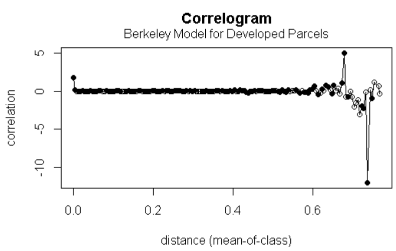
(a) Standard QQ plot showing that residuals are close to normally distributed.



(b) Bubble plot. The size of the dots is proportional to the value of the residuals, and different colors are used for positive and negative residuals.



(c) Spatial correlogram. It plots Moran's Index as a function of distance (1 in x-axis = 60 miles).



(d) Experimental variogram.It groups distances into classes and displays the average semivariance among each class (1 in x-axis = 60 miles).

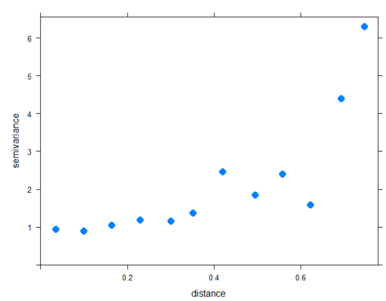
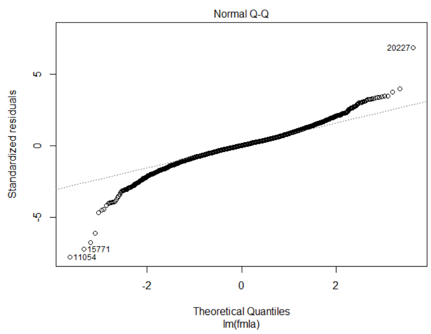
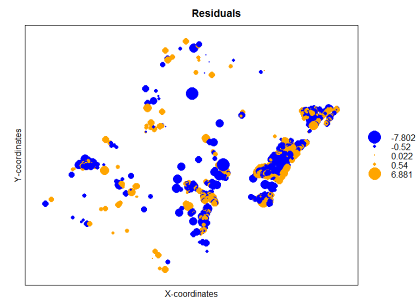


Figure D.5: Diagnostic tools used to visually assess existence of spatial dependence in the error term of the hedonic model of residential parcels in Berkeley county.

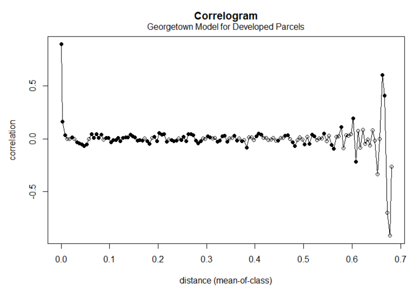
(a) Standard QQ plot showing that residuals are close to normally distributed.



(b) Bubble plot. The size of the dots is proportional to the value of the residuals, and different colors are used for positive and negative residuals.



(c) Spatial correlogram. It plots Moran's Index as a function of distance (1 in x-axis = 60 miles).



(d) Experimental variogram.It groups distances into classes and displays the average semi-variance among each class (1 in x-axis = 60 miles).

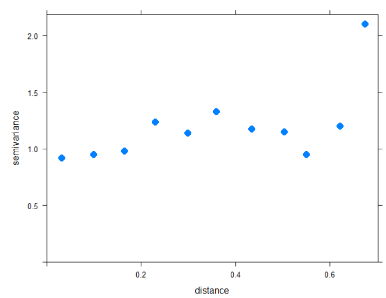


Figure D.6: Diagnostic tools used to visually assess existence of spatial dependence in the error term of the hedonic model of residential parcels in Georgetown county.

E Sample selection bias in hedonic equations

Developed land is not assigned across the landscape at random. Moreover, it is likely that unobserved characteristics inherent to a parcel, are correlated with both its current–use and its value. For example, an owner of agricultural land that perceives risk of flood hazard to be high, may purchase protective insurance or incur further investments on the property that make it more valuable in the eyes of developers. Thus, this particular parcel, is more likely to be developed and sell for higher prices than other parcels.

The hedonic price model for developed land is estimated on data from parcels that are already developed. Thus, if there are unobserved factors that influence both the decision to develop a parcel and its subsequent market value, the estimated coefficients will be biased. For instance, if residential parcels tend to fetch higher prices than agricultural prices, and if parcels that are closer to the beach are more likely to develop,then, the effect of proximity to the beach on sales prices will seem larger than its true effect. This is problematic because the purpose of estimating the residential hedonic models is to infer unknown sales values for lands that are currently in agriculture, but if the sample of developed parcels is not representative of the general parcel population, the estimated coefficients are biased and lead to invalid inference.

The problem of sample selection bias calls for corrective measures. Other works in the literature have opted to proceed to develop land use change models without taking corrective measures (Irwin and Wrenn, 2014; Newburn et al., 2004-2006; Bockstael, 1996). In fact, few works empirically address the issue of sample selection bias, even though the literature clearly recognizes the problem. A theoretical justification for abstracting away from the sample selection problem, is the idea that a hedonic model of residential land value may be unbiased if developers' ignorance of omitted variables that drive the selection problem, matches the ignorance of the researcher (Bockstael, 1996). This appendix includes findings from the exploratory analysis of instruments for a selection bias correction. To illustrate the trivial effect of including or excluding weak instruments from estimation, results from hedonic regressions that correct for sample selection bias using the proposed weak instruments and the Heckman methodology are presented. In general, they are note different in direction, magnitude, or significance from results kept for the main estimation.

To estimate valid (i.e., unbiased) coefficients in a hedonic model for developed parcels, the corrective procedure proposed by Heckman (1979) can be used. There are two steps in this procedure. First, a selection equation describing the probability that a parcel is developed is estimated using a standard probit model using maximum likelihood estimation (MLE). Second, using the results from the probit estimation, an Inverse Mill's Ratio variable is formed and included as one of the explanatory variables in the hedonic price model for developed parcels.[27]

[27]The Inverse Mill's Ratio variable (IMR) is the ratio is the ratio of the probability density function to the cumulative distribution function of a distribution. It is probability that a landowner decides to develop his land over

The selection model can be formalized as:

$$s_i y_i = X_i\beta + \upsilon, \tag{E.1}$$

where s_i is a selection indicator, y_i is a latent variable reflecting the sales price of residential properties, $X_i\beta$ is the deterministic component that defines sales prices, and υ is a random component. The latent variable y is not observed for every parcel, instead, it is only observed for parcels that are already developed. Therefore, the selection indicator s_i is equal to 1 for developed parcels, and 0 for undeveloped parcels.

This type of sample selection is called incidental truncation. It receives that name because the rule determining whether or not sales prices of residential properties are observed, does not directly depend on the actual sales price. The usual approach to incidental truncation is to add an explicit selection equation to the model shown above. Explicitly, the selection model becomes:

$$y = X\beta + \upsilon, \mathrm{E}(\upsilon|X) = 0 \tag{E.2}$$

$$s = \mathbf{1}[Z\gamma + \nu], \tag{E.3}$$

where $s = 1$ if y is observed, and zero otherwise. In turn, Z is a matrix of covariates containing parcel characteristics that are exogenous to residential property values and of which X, in equation (E.2), is a strict subset.

Under Heckman's procedure, to properly identify the effects of the IMR variable, Z must include an instrumental variable that is sufficiently correlated with the selection decision, shown in equation (E.3), but not with the sales price of residential properties, shown in equation (E.2). This instrument, Z_j, is referred to as the exclusion restriction variable (i.e., the element in Z that is not in X).[28]

To implement Heckman's sample selection corrective method to my empirical application, this study uses as the exclusion restriction variables multiple factors that affect agricultural land value but not residential land value. This rationale for instrument choice (i.e., using variables that affect agricultural value of a parcel and therefore its probability of development but are not related to the residential value of land) is found in the literature.

Explicitly, as instruments, measures of soil quality and proximity to agricultural processing facilities are used. Intuitively, soils that are better for agriculture would discourage conversion of agricultural land but do

the cumulative probability of the landowner's decision. The IMR's coefficient is interpreted as the fraction of the covariance between the decision to develop a parcel and the net benefit from developing that parcel relative to the variation in the decision to develop it.

[28] This is required to avoid problems with large standard errors when estimating (E.2). Because the IMR is a linear function of X, having every Z_j in X would result in multicollinearity. A thorough exposition of this method can be found in Wooldridge's Introductory Econometrics.

not impact the developed value of a parcel (to my knowledge, there are no studies suggesting that different soil qualities, at least as defined by the USDA nomenclature, facilitate or hinder construction). Similarly, proximity to processing facilities makes agricultural production more attractive by reducing costs.

Table E.1 shows the findings from the exploratory analysis to determine the exclusion variables. Despite the theoretical strength behind the logic used to select these instruments, a close examination of table E.1 offers weak support for the choice of instruments.

The condition for a valid instrument is that is is significantly related to the selection equation but has no direct association with sales prices. For all counties, it is found that none of the instruments is empirically sound as most of the time (that is, with the more parsimonious model and with a model that is highly specified), they are significantly related to probability of development and to sales prices.

Using weak instruments is likely to produce inconsistent estimates (Bound, Jaeger and Baker, 1995). In addition, even with large data sets, using weak instruments can lead to estimates that are biased in the same direction as OLS estimates, with the magnitude of the bias approaching that of OLS. This is evident in the minimal difference between the results from hedonic equations that use a corrective step and the hedonic equations used to estimate the results shown in the main text of this study.

Table E.2 shows the results from the probit estimations to correct for sample selection bias in the hedonic models for residential properties. The probit results are used to construct an IMR that is then used as an explanatory variable in the hedonic equation.

The results from the sample selection models generally conform to prior expectations. Most important for the sample selection results, is the fact that the coefficients of most exclusion variables are significant and have the appropriate negative sign for the models of Charleston and Berkeley counties. For Georgetown, some have positive signs and some are not significant.

The factors that are more important in determining probability of development across all three counties are parcel size, elevation, proximity to major roads and beaches with public access, as well as spatial measures of landscape composition. Overall, relative to undeveloped parcels, developed lands in all three counties are smaller and are located in higher elevations. Also, parcels more heavily surrounded by by developed covers are more likely to develop.

In Charleston county, developed parcels tend to be more proximal to major roads, and parcels that are closer from public beach access point are less likely to develop. In Berkeley and Georgetown counties, being further from Charleston city center is positively associated with the probability of development, as is being further away from major roads. This discrepancy may indicate that open space and distance from noise and pollution corridors (e.g., major roads) are more valuable features for landowners in Berkeley and Georgetown counties than in Charleston. Nevertheless, this effect is small.

In Berkley and Georgetown, parcels that are closer to the beach are more likely to be developed. Also, the further a parcel is from a hurricane evacuation route, the more likely it is to be developed. This result may provide early indication that landowners and developers take into consideration risks of flooding when choosing to develop a parcel, a finding supported by recent empirical evidence (McCoy and Zhao, 2018). In their research, McCoy and Zhao use data on capital investment projects in homes located in flood risk areas. They find evidence that recent flooding may increase perceived risks and investment in buildings.

Finally, results regarding the effect of landscape composition on development offer important insight for the design of land use policies aimed at slowing down development in areas vulnerable to flood risks. First, wetland covers are generally significantly related to the probability that a parcel develops, although the effect is small. In Georgetown county the effect is negative, which could be related to conversion costs that can include legal costs of obtaining 404 permits (i.e., permits to dredge or fill resources that may include wetlands) or expenses to cover required compensatory mitigation (these legislative costs are described in detail in section 5).

Second, the prevalence of developed covers (which include a mixture of construction materials and vegetation) in neighboring spaces positively and significantly influences the probability that parcels develop in all three counties. That wetland cover sometimes significantly decreases the likelihood of parcels developing suggests that implementing policies that encourage the conservation of wetland habitats may be an effective tool for managing urban growth in coastal counties like Charleston and Georgetown. In turn, the positive effect of developed covers on conversion suggests that designing policies that discourage dense development (for instance a tax on high intensity developed covers) may also slow down urban growth in counties similar to those studied here.

Table E.3 shows the results from the estimation of hedonic equations for residential properties. Importantly for the empirical validity of this study, the coefficient on the IMR variable derived from the sample selection model is positive and significant in the estimation of models for Charleston and Berkeley counties.

Across counties, age of residential buildings in the parcel, square footage of the structure, and assessed value of the structure show a significant quadratic relationship with developed sales prices. Younger structures are associated with higher prices as are structures with higher assessed values. However, building size has different effects across counties. In Charleston county, building square footage is positively related to sales price while the opposite is true for the other counties, although the effect is not significant in Berkeley.

Parcel elevation is negatively related to sales prices as is the dummy variable indicating whether a river runs through or is adjacent to the parcel. The landscape variables measuring crop and pasture cover as well as percent of forested area are also negatively related to sales prices. Whether a river is present and the proportion of crop and forest cover in a parcel are variables that may play a role in conversion costs.

An interesting finding is the effect of flood hazards on sales prices (as indicated by a binary measure of whether parcels are located within the 100–year flood plain). In Charleston county, being within the flood plain is negatively and significantly related with sales prices. This negative effect provide indication that residential land markets may react differently to flood risks than agricultural land markets due to fundamental reasons that have not yet been studied. In Berkeley and Georgetown counties, the effect of flood hazards on property values is not significant, suggesting property markets in these two counties may be fundamentally different from that of Charleston, favoring the use of separate land use change models in the second stage of the empirical analysis.

In general, proximity effects are negligible in terms of magnitude. However, they point towards the existence of underlying differences in land markets across counties. In Charleston county, proximity to downtown Charleston, major roads, access points to public beaches, the coastline, and the nearest river positively affect sales prices. In Berkeley county, almost all these relationships show the opposite direction. In Georgetown county, the effect of proximity variables resemble those from Charleston, except for proximity to downtown Charleston which is positively related to sales prices.

The effects on sales prices from landscape composition variables are more consistent across counties. Increases in the density of developed covers surrounding parcels has a positive effect on sales prices, an intuitive effect of urbanization that can work by lowering distribution costs of public utilities, for instance. Finally, increases in the density of wetland covers surrounding parcels has a positive effect on sales prices in Charleston, a negative effect in Berkeley, and an insignificant effect in Georgetown. This outcome is of interest for designing land use policies aimed at conserving natural infrastructure that provides flood protection services. It suggests there may be larger gains from imposing such a policy in Charleston than in Berkeley or Georgetown counties.

Table E.1: Correlation between exclusion variable and selection and value processes.

Exclusion variable	Process: $s = \mathbf{1}[Z\gamma + \nu]$ Charleston	Berkeley	Georgetown	$y = X\beta + \upsilon$ Charleston	Berkeley	Georgetown	Controls^{+}
Soil quality^{++}	-0.262	-0.616**	1.286***	-0.078	-0.330***	-0.759**	YES
	(0.293)	(0.240)	(0.457)	(0.107)	(0.089)	(0.372)	
	-0.168	0.351**	1.158***	-1.995***	-0.310**	-1.807***	NO
	(0.254)	(0.165)	(0.307)	(0.220)	(0.123)	(0.321)	
Dist. to mill (mi)	-0.136*	-0.044**	0.127***	-0.172***	-0.044***	0.020	YES
	(0.080)	(0.021)	(0.030)	(0.026)	(0.011)	(0.028)	
	-0.076***	-0.027**	0.072***	-0.022***	-0.043***	0.103***	NO
	(0.006)	(0.011)	(0.010)	(0.006)	(0.008)	(0.010)	
Soil quality * Dist. mill	0.024	0.043	-0.018	0.006	0.017	0.011	YES
	(0.021)	(0.031)	(0.033)	(0.010)	(0.014)	(0.026)	
	-0.004	-0.049**	-0.131***	0.116***	-0.063***	-0.110***	NO
	(0.019)	(0.022)	(0.024)	(0.021)	(0.017)	(0.023)	

Significance codes: *p<0.1; **p<0.05; ***p<0.01
The errors in parenthesis are robust standard errors.
$^{+}$ "NO" means only the exclusion variables were included. "YES" means the model is as shown in table E.2
$^{++}$ Soil quality is constructed as a weighted sum of soil quality. Higher scores are better for agriculture.

Table E.2: Sample selection correction model.

	Dependent variable: Prob(Developed = 1)		
	Charleston	Berkeley	Georgetown
Constant	1.951***	-0.490	-11.654**
	(0.745)	(0.693)	(4.694)
Soil quality	-0.262	-0.616**	1.286***
(exclusion variable)	(0.293)	(0.240)	(0.457)
Dist. to mill (mi)	-0.136*	-0.044**	0.127***
(exclusion variable)	(0.080)	(0.021)	(0.030)
Soil quality * Dist. to mill	0.024	0.043	-0.018
(exclusion variable)	(0.021)	(0.031)	(0.033)
Acreage	-0.025**	-0.21***	-0.001*
	(0.01)	(0.048)	(0.001)
Acreageˆ2	0.00003***	0.013**	$0.005e^{-4}$
	(0.00001)	(0.006)	($0.003e^{-4}$)
Elevation (m)	0.065***	0.013**	0.029**
	(0.015)	(0.006)	(0.011)
River (=1)	0.061	-0.027	-0.022
	(0.081)	(0.061)	(0.081)
Dist. to Charleston (mi)	-0.052	0.047	0.282**
	(0.08)	(0.08)	(0.136)
Dist. to Charlestonˆ2 (mi)	0.008***	-0.003*	-0.002*
	(0.003)	(0.002)	(0.001)
Dist. to road (mi)	-0.381***	0.192***	0.323***
	(0.063)	(0.06)	(0.078)
Dist. to roadˆ2 (mi)	0.057***	-0.061***	-0.107***
	(0.012)	(0.016)	(0.028)
Dist. to beach (mi)	0.407***	-0.232***	-0.098***
	(0.058)	(0.086)	(0.036)
Dist. to beachˆ2 (mi)	-0.279***	-0.312***	-0.084**
	(0.065)	(0.104)	(0.038)
Dist. to evacuation route (mi)	X	0.185***	0.166***
		(0.043)	(0.041)
Surrounding wetland cover	0.002***	0.004***	-0.003**
	(0.001)	(0.001)	(0.002)

Continued on next page

Table E.2 (continued).

	Charleston	Berkeley	Georgetown
Surrounding developed cover^{+}	0.01***	0.01***	0.005***
	(0.001)	(0.001)	(0.001)
Surrounding pasture/crop cover	0.003*	0.0001	0.002
	(0.002)	(0.002)	(0.003)
Surrounding forest cover	-0.001	-0.005***	-0.002
	(0.001)	(0.001)	(0.002)
Year fixed effects	YES	YES	YES
Tract fixed effects	YES	YES	YES
Block group fixed effects	YES	NO	NO
Observations	17,111	16,793	6,065
Log Likelihood	-4,768.4	-6,516.6	-3,356.5
Akaike Inf. Crit.	9,880.8	13,163.2	6,792.9

The above results correspond to binary probit models.
Coefficients do not show marginal effects.
$^{+}$Includes open space-, low-, medium-, and high-intensity development.
Significance codes: *p<0.1; **p<0.05; ***p<0.01
The errors in parenthesis are robust standard errors.

Table E.3: Corrected hedonic model for residential parcels.

	Dependent variable: $log(\frac{Price}{Acre})$		
	Charleston	Berkeley	Georgetown
Constant	16.136***	13.075***	-11.73*
	(0.241)	(0.415)	(6.125)
IMR	0.0024	0.025$^{\cdot}$	-0.008
	(0.017)	(0.013)	(0.024)
Age	-0.03***	-0.059***	-0.032***
	(0.001)	(0.003)	(0.003)
Ageˆ2	0.00032***	0.001***	0.0003***
	(0.00001)	(0.0001)	(0.00005)
SQFT	0.0002***	-0.00001	-0.0002***
	(0.00003)	(0.00003)	(0.00003)
SQFTˆ2	-0.003e^{-5***}	0.006e^{-6*}	0.003e^{-6***}
	(0.003e^{-6})	(0.003e^{-6})	(0.007e^{-7})
Building value	0.00001***	0.003e^{-4***}	0.003e^{-3***}
	(0.006e^{-4})	(0.006e^{-5})	(0.002e^{-4})
Building valueˆ2	-0.006e^{-11***}	-0.026e^{-12}	-0.001e^{-9***}
	(0.001e^{-11})	(0.005e^{-11})	(0.002e^{-10})
100-year flood plain (=1)	-0.062***	0.038	-0.007
	(0.013)	(0.024)	(0.033)
Elevation (m)	-0.017***	0.002	-0.106***
	(0.007)	(0.003)	(0.010)
River (=1)	-0.401***	-0.224***	-0.284***
	(0.039)	(0.034)	(0.069)
Dist. to Charleston (mi)	-0.063***	0.078**	0.639***
	(0.039)	(0.033)	(0.174)
Dist. to Charlestonˆ2 (mi)	0.005	-0.011***	-0.004***
	(0.002)	(0.001)	(0.001)
Dist. to road (mi)	-0.044	0.138***	-0.073
	(0.029)	(0.031)	(0.058)
Dist. to roadˆ2(mi)	0.007	-0.025**	0.032
	(0.007)	(0.010)	(0.020)
Dist. to beach (mi)	-0.138***	0.014	-0.065

Continued on next page

Table E.3 (continued).

	Charleston	Berkeley	Georgetown
	(0.037)	(0.049)	(0.046)
Dist. to beachˆ2 (mi)	0.004	0.007***	0.010***
	(0.003)	(0.001)	(0.002)
Dist. to coastline (mi)	-0.193**	X	-0.153***
	(0.044)		(0.042)
Dist. to coastlineˆ2 (mi)	0.012**	X	-0.007**
	(0.004)		(0.003)
Dist. to river (mi)	-0.406***	X	-0.287
	(0.134)		(0.222)
Dist. to riverˆ2 (mi)	0.537**	X	0.847**
	(0.235)		(0.391)
Surrounding wetland cover	0.003***	-0.001***	0.001
	(0.0004)	(0.0003)	(0.001)
Surrounding developed cover+	0.005***	0.002***	0.006***
	(0.0003)	(0.0002)	(0.001)
Surrounding pasture/crop cover	-0.002***	-0.019***	-0.003
	(0.001)	(0.001)	(0.003)
Surrounding forest cover	-0.0035***	-0.0018***	-0.0028**
	(0.0004)	(0.0003)	(0.001)
Year fixed effects	YES	YES	YES
Tract fixed effects	YES	YES	YES
Block group fixed effects	YES	NO	NO
Observations	14,429	12,315	3,627
R^2	0.769	0.695	0.709
Adjusted R^2	0.766	0.694	0.705
Residual Std. Error	0.51	0.45	0.575
	(df = 14254)	(df = 12247)	(df = 3581)
F Statistic	273***	423***	194***
	(df = 174; 14254)	(df = 67; 12247)	(df = 45; 3581)

Building values are assessed values and are adjusted to inflation to $2016.
+Includes open space-, low-, medium-, and high-intensity development.
Significance codes: *p<0.1; **p<0.05; ***p<0.01
The errors in parenthesis are robust standard errors.

F Derivation of development probabilities

As presented in section 3, a hazard function is also the derivative of a survivor function S_t. A survivor function is the probability that an event (e.g., development of a parcel) does not occur until after period t, and it is equivalent to $1 - Pr(T \leq t)$, or $Pr(T > t)$, where $Pr(T \leq t)$ is the probability that the duration is less than t, or the cumulative distribution function of T, or of how long a parcel remains undeveloped. The survival function is also related to the cumulative hazard or integrated hazard function, Λ_t, which shows the expected number of events (in this case developments) for a given set of covariates at a particular time and is equal to $-\ln S_t$, or equivalently, $S_t = \exp(-\Lambda_t)$.

To derive development probabilities in this study, the hazard model estimates are manipulated to compute survival probabilities for each undeveloped parcel in the sample and simply set the probability of development to be the probability that a parcel does not survive in an undeveloped state up to time t, or that it is converted within time period t (i.e., $\delta_{it} = 1 - S_{it}$).

With the proportional hazard assumptions, the survival function at time t is:

$$\begin{aligned} S_t &= \exp\Big(- H_{0t} \exp(X^{'}\beta)\Big) \\ &= \exp(-H_{0t})^{\exp(X^{'}\beta)}, \end{aligned} \tag{F.1}$$

where H_{0t} is the baseline cumulative hazard function at time t (an intercept–like term that describes the risk for parcels with covariates X equal to zero); X is a set of covariates; and β is a vector of coefficients that measure the impact of covariates.

The R function basehaz() provides the estimated hazard function H_{0t}, while the coxph() function yields coefficient estimates β for a given sample.

Because of the proportional nature of the proportional hazard model, to obtain survival curves of the 2 groups defined by their values of a particular covariate, say the developed value of land, the $\widehat{H_{0t}}$ can be raised to the power of the estimated coefficients for that covariate. For example, for $\ln \frac{P^D}{Acre} = \{\$10,000, \$9,000\}$, the survival curves are:

$$S_{t|\$10,000} = \exp(-\widehat{H_{0t}})^{\exp(\hat{\beta}\cdot 10,000)} \tag{F.2}$$

$$S_{t|\$9,000} = \exp(-\widehat{H_{0t}})^{\exp(\hat{\beta}\cdot 9,000)} \tag{F.3}$$

Presumably, the group of parcels with higher developed values has a lower survival curve than the group with lower developed values.

Using the predict() function, the cumulative hazard rate is derived for a given set of covariates at a particular time that is associated with the $\hat{\beta}$ that were estimated with the coxph() function. The prediction commonad is specified so that the prediction of the cumulative hazard rate incorporates the baseline hazard. Thus, the predicted cumulative hazard represents the absolute hazard, instead of the hazard relative to the sample average. Moreover, the coxph() command is specified to use the baseline hazard for time 5. Explicitly, the probability that a parcel survives development for five years is predicted using the following statements:

$$\begin{aligned}\hat{S}_5 &= \exp\big(-\widehat{H_{05}}\exp(X'\hat{\beta})\big)\\ &= \exp\Big(\text{-predict}\big(\text{coxph}(h_t = h_{0t}\exp[\beta_D \ln\frac{P^D}{Acre} + \beta_U \ln\frac{P^U}{Acre}], \text{data}_{t=5})\big)\Big).\end{aligned} \tag{F.4}$$

The predicted development probabilities are:

$$\hat{\delta}_5 = 1 - \hat{S}_5. \tag{F.5}$$

To extract development probabilities in the simulation exercise, values of land that have been modified to reflect the policy rules in each scenario are inputed into the calculation.

Specifically, the coefficients resulting from estimating the proportional hazards model on the original data are interacted with parcel characteristics that have been altered according to the policy. The alternative survival probabilities are:

$$\begin{aligned}\tilde{S}_5 &= \exp(-\widehat{H_{05}}\exp(\tilde{X}'\hat{\beta}))\\ &= \exp\Big(\text{-predict}\big(\text{coxph(model), data}_{X=\tilde{X}}\big)\Big).\end{aligned} \tag{F.6}$$

where the model is $h_t = h_{0t}\exp[\beta_D \ln\frac{P^D}{Acre} + \beta_U \ln\frac{P^U}{Acre}]$, and is estimated using original data where time has been set to 5, or $\text{data}_{t=5}$.

There are concerns over using the predicted baseline hazard to project out of sample development patterns (i.e., to use it for future time periods not observed in the data used to estimate the econometric model). Strictly speaking, the estimated proportional hazards model with time invariant regressors X and baseline hazard estimate $\widehat{H_{05}}$, is only suitable for predicting counterfactual development probabilities within the time period observed in the data, not for predicting future development probabilities for time periods not observed in the data.

The basehaz() function estimates the cumulative baseline hazard and not the baseline hazard rate itself. When plotted, there is a time trend in the cumulative baseline hazard. This is shown in figure G.1. In the simulation of scenarios five years into the future, the cumulative baseline hazard for $t = 5$, also shown in figure G.1, is used.

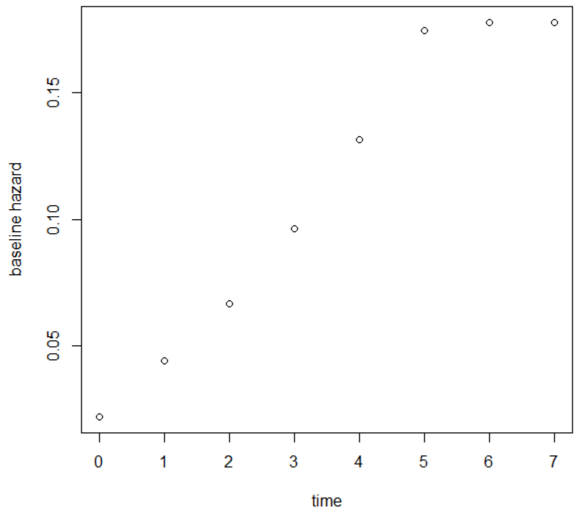


Figure F.1: Cumulative baseline hazard using basehaz() function on original data.

G Development of the flood risk index

A central interest of this project, is to examine how changes in impervious and wetland covers influence a parcel's vulnerability to flood events. The general understanding in the natural resources literature is that urbanization changes a basin's response to precipitation, therefore changing flood frequency characteristics of urban streams. The most common effects of urbanization on stream characteristics are reduced infiltration of precipitation into the soils and more rapid runoff (or decreased lag time), which substantially increase runoff volume and flood frequency (Feaster, Gotvald, and Weaver, 2011).

Factors that affect the average capacity of a parcel to provide flood preventive services include elevation, proximity to water bodies, as well as how impervious surfaces, wetland habitats, and tree cover are spatially distributed in and around the parcel. Therefore, this study uses a measure of flood risk that is a function of landscape characteristics to calculate the economic benefits of imposing different land use policies in the Lowcountry.[29]

Currently, there are no available flood maps or flood risk products that can be used for the purpose of this study. As of 2018, digital flood risk maps only exist for three counties in South Carolina, and only one of them is in my study area (Berkeley). FEMA manages flood maps for about 22,000 communities across the U.S. Almost two–thirds were officially updated more than five years ago. Some maps have been in place for more than 40 years. In addition, even updated maps may not be suitable for analysis at the parcel level.[30] South Carolina has not been hit by a major hurricane storm since Hurricane Hugo in 1989. Thus, recent mapping techniques have not been applied to this study area to determine inundation after a storm. In addition, because of the cloud cover that accompanies storms, satellite images are difficult to use for measuring inundation and flooding during and immediately after a storm. Hence, due to the lack of appropriate measures, a

[29] Recall that benefits are defined as the sum of property damages from major floods that are avoided because of a particular land use policy, such as a policy discouraging land owners from adding hard surfaces to their land or a policy encouraging them to conserve wetlands that are present in their lots.

[30] The most recent recent flood model being used by FEMA, the HAZUS–MH flood model. The model defines annual probability of coastal inundation using topographic elevation data. Other inputs include dune erosion, wave effects, and regional riverine water discharge. There are available tools, other than FEMA's flood plain maps, to model inundation, including NOAA's SLAMM and SLOSH models, which project future scenarios under sea level rise and changing storm surge conditions, respectively. However, these models appear to be better suited for analysis at a scale larger than the parcel–level. These models are also criticized for lacking structural strength, i.e., they are mostly statistical extrapolations rather than fundamental models of underlying natural process driving observed events.

flood risk index is constructed based on the foundations of existing inundation models using data from recent large storm events in South Carolina.

The flood risk index is developed to reflect the dynamics in the natural composition of the landscape so that a parcel that is 5% covered by wetland habitat and 95% covered by impervious surfaces shows a different risk of flooding than a parcel that is 5% wetland, 25% forest, and 75% by impervious surfaces or a parcel that is 25% wetland, 25% forest, and 50% impervious surfaces. Specifically, separate probabilistic models are developed using the data on Hurricanes Joaquin (2015) and Irma (2017) as the dependent variable, and the probabilistic measure of propensity to flooding is modeled as a function of topographic characteristics, like elevation and proximity to water bodies, and the land cover composition of a parcel and its surroundings.[31]

The model is defined as:

$$\begin{aligned} Pr(\text{Flood}_{iht}) = f(&\text{elevation, proximity to a river, proximity to the coastline,} \\ &\text{parcel land cover, land cover in surrounding area}), \end{aligned} \tag{G.1}$$

where $\Pr(\text{Flood}_{iht})$ is the probability that parcel i gets inundated during Hurricane h at time t. Variables characterizing parcel land cover are share of the parcel's area under high–intensity development covers, forest covers, and wetland covers. In turn, variables representing land cover in the area surrounding the parcel, are shares of high–intensity developed covers, a mix of developed and vegetation covers, forest covers, and wetland covers. The model is estimated using a probit method on two separate datasets, one corresponding to Hurricane Joaquin and one corresponding to Hurricane Irma.

G.1 Data on property damage from major floods

For the average property damage (i.e., $\overline{\text{Damage}}$), three different estimates of average property damage are exlplored to develop an idea of confidence around existing measures but only one is used in the final calculation of costs and benefits (the role of this measure in the simulations analysis is further discussed in sections 4 and 5). The explored estimates come from three sources: data from HAZUS®–MH Flood Module, a recent estimate of average flood damages from Hurricane Sandy

[31] For these measures, a buffer zone is defined. The zone has a 0.25 miles radius from the parcel's centroid (an area equivalent to 0.2 squared miles).

to residential properties in three coastal counties of New Jersey, and the 2016 average National Flood Insurance Program (NFIP) payment per claim in South Carolina. In 2016 dollars, the three different values used to calibrate the average residential property damage from a major flood are: $273, $22,672, and $24,776. Because the simulations procedure uses data from the National Flood Insurance Program (NFIP) to approximate flood damages, it is the only data discussed in the text.

South Carolina ranks among the top states in the country in terms of FEMA payments made for insured losses through the NFIP. In 2016 alone, South Carolina was the third state with the largest NFIP claim payments, just behind Louisiana and Texas (payments in 2016 were $139.8 million in SC, $696 million in TX, and $805 million in LA), and as of July 2017, South Carolina was the sixth state with most NFIP policies in force. Between July 2016 and July 2017 (before Hurricanes Harvey, Irma, and Maria), nationally, the average payment per claim was $31,593, but in SC it was much higher at $38,217.[32] In 2016, there were 5,643 NFIP claims in SC and the amount of claim payments was $139.81 million. Thus, the average NFIP payment per claim in SC was $24,776.[33] In calculating the benefits of implementing land use policies, this measure is used to approximate property costs from flooding.[34]

G.2 Flood risk data

For the parcel–specific measure of flood vulnerability, the analysis uses predicted probabilities of flooding derived from an inundation model that is developed and estimated using data from the most recent large storm events in South Carolina: Hurricane Irma in 2017 and Hurricane Joaquin in 2015.[35] This aspect of the research was developed in partnership with scientists at the South East Climate Adaptation Science Center.[36]

[32]Between 1978 and June 2017, FEMA payments in SC were over $748 million. Approximately $192 million were sent to Charleston county (roughly 25% of the state's losses). In turn, FEMA paid more than $105 million to Berkeley (14% of the state's claims) and $113 thousand to Georgetown (III, 2017).

[33]For reference, consider that the average NFIP payment nationwide in 2016 was 62,247. However, in the last decade, average claim payouts vary widely from year to year. From $25,133 in 2009 to $83,198 in 2005 (the year of Hurricane Katrina).

[34]NFIP statistics on total claim payments for the 2017 fiscal year at the state level (when Hurricane Irma occurred) are not yet publicly available as some payouts are still being processed. Thus, data on 2016 claim payments are used. For reference, in 2017, the year of Hurricanes Harvey, Irma, and Maria, the average claim payment nationwide was $91,735. That year there were 83,779 NFIP policies in the three–county study area (67,829 in Charleston, 7,493 in Berkeley, and 8,457 in Georgetown) covering a total of $23,305,207,700 (Insurance Information Institute, 2018).

[35]Data from 2018 Hurricane Florence are not available and prior to 2015, the only other large storm affecting the region was Hurricane Hugo in 1989.

[36]The scientists that informed the development of this FRI are Mitchell Eaton, a research ecologist at the SECASC and the US DOI, and Simeon Yurek, a wetland ecologist at the SECASC, USGS Wetland and Aquatic Research

To develop the flood risk variable for each parcel in the sample, information on actual flooding following two severe storm events is combined. Specifically, flood inundation maps created by the US Geological Survey (USGS) to model the incidence of Hurricane Joaquin in 2015, and real time simulation models produced by the Coastal Emergency Risks Assessment (CERA) group to reproduce the effects or Hurricane Irma in 2017 are used. All these data are publicly available and can be viewed and managed using GIS software.

In the paragraphs that follow,data used to develop a flood risk index (FRI) and then present the methodology followed to derived said measure but before doing so is described. It is worth noting that the derived measure of flood risk is not directly included in the hedonic models. This is due to theoretical and econometric considerations. From the economic theory standpoint, the FRI variable would be included as determinant of the price function if landowners and investors followed it to develop expectations over land price. However, evidence from literature in psychology and behavioral economics does not always support the existence of this response, and yet it is well established that flood risk plays a significant role in determining insurance premiums and other costs that landowners incur.[37] Thus, instead of explicitly including the flood risk as a explanatory variable, a binary variable showing when a parcel is within the flood plain is included in the price equations. Additionally, land cover characteristics that are included as regressors in the flood risk equation are also included in the hedonic price function. Thus, to the extent that these land cover features reflect changes in flood risk, the hedonic equations do take into consideration landowners' preferences for changes in flood risk. The econometric reason that further justifies excluding the derived flood risk measure from the hedonic estimation is that the FRI variable is generated and not obtained from an external sources, which poses questions of measurement error.[38]

G.2.1 USGS inundation maps

During October 1 to 5, 2015, heavy rainfall related to Hurricane Joaquin occurred across South Carolina. The storm caused major flooding in the central and coastal parts of the State. The maximum recorded rainfall was almost 27 inches in Charleston county, 22 in Berkeley, and 20.75 in

Center and the South Atlantic Landscape Conservation Cooperative.

[37]See Tversky and Kahneman (1974), and Kunreuther and Pauly (2006).

[38]As econometric theory indicates, measurement error in the explanatory variable biases the estimates towards zero and can correlation between explanatory variables and the error term can arise, posing problems of endogeneity (Hayashi, 2000).

Georgetown (USGS, 2016). Using high–water marks and streamflow measurements, USGS personnel created 20 inundation maps of 12 urban communities. The maps modeled boundary and water depth of the inundations. The digital maps include four inundation zones in coastal Charleston and Georgetown counties: Coastal Charleston, Coastal Georgetown, Black River community, and the Ashley River community. According to these digital images, a total of 4,472 parcels in Charleston and Georgetown counties are shown to have flooded with Joaquin, and the average rainfall was 2.2 inches. Figure G.1 displays the location of these four inundation maps.

G.2.2 CERA simulations of storm path

The Coastal Emergency Risks Assessment (CERA) group develops real–time forecasting of storm models to provide guidance about storm surge and waves for the Atlantic Coast. The information produced by CERA is based on two models: the Advanced Circulation and Storm Surge model (ADCIRC) for surge and Simulating Waves Nearshore (SWAN) for waves. The CERA team provides all layers in shapefile format, which include comprehensive information of particular storm events. This study uses CERA's simulations of Hurricane Irma to measure inundation in the study area.

Figure G.2 shows CERA's map of maximum water height in feet above mean sea level (MSL) of Hurricane Irma in the three counties of the study area.[39] Of the 237,754 parcels in the three counties, 7 percent flooded with Irma and the average water depth was 6.9 feet above MSL, which is just higher than the typical high tide in the region.[40]

G.3 Estimation results

An additional step was taken before setting up the probabilistic model using the maps created by the USGS to identify inundated areas during Joaquin. Because these maps do not constitute exhaustive maps of inundated areas (i.e., there are zones in the study area that flooded and were not mapped), a 0.25 mile radius buffer zone around the four inundation maps was defined. Only parcels within

[39] Height above mean sea level is the elevation (on the ground) or altitude (in the air) of an object, relative to the average sea level. A common measure of average sea level is the midpoint between a mean low and mean high tide at a particular location. A typical high tide at the Cooper River station (ID 8665530) is 6.76 ft above MSL.

[40] As a point of reference, consider that a recent study by Loerzel et al. (2017) calculates that approximately 7.6 percent of the parcels in their study area flooded during Hurricane Sandy, and the mean flood depth in feet above MSL among flooded parcels was 3.5. This information seems to give support to the inferences from inspecting Irma's data.

the buffered zone were considered in estimating the probabilistic model. Thus, parcels within the buffered area and inside the inundation map are recorded as flooded while parcels in the buffered area but outside the inundation map are considered as not flooded. Figures G.3 and G.4 illustrate the process of designating parcels as flooded or not during each event.

Results from the probit estimations are presented in table G.1. As expected, factors that increase probability of flooding are lower elevation, proximity to water bodies, and prevalence of impervious surfaces in the area surrounding a parcel. Alternatively, the factors that decrease probability of flooding are forest cover and wetland cover in the area surrounding a parcel. The significance of the interaction terms indicate that absolute area of impervious surfaces, and not just relative prevalence, matter in determining the probability of flood, and also that surrounding hard covers increase the probability of flooding slightly more heavily for parcels that are closer to a river.

These results are consistent with findings in the literature. In a study of streamflow and urbanization, Feaster, Gotvald and Weaver (2011) find that a one percent increase in impervious surface cover increases streamflow during a 100–year flood by 0.3 to 3.9%. In turn, ecosystem service valuation studies generally find that wetland covers reduce hurricane damages by attenuating the effect of flooding.[41]

Interestingly, the relative presence of impervious surface in the parcel does not significantly affect the likelihood of it flooding during a hurricane. Only the composition of the area surrounding the parcel seems to matter in this regard. According to the estimation, a one percent increase in total surrounding area covered by impervious surfaces increases the likelihood of flooding from 7.14 to 26.5%. Finally, it is surprising to find that prevalence of general developed covers (i.e., a mix of construction materials and vegetation) *negatively* influences the probability of flood when the model is estimated using data for Hurricane Irma. This suggests that urbanization around a parcel can decrease its chances of flooding, which can occur is impervious surfaces speed up the removal of flood flows from upper parts of stream (Feaster, Gotvald, and Weaver, 2011).

[41]Estimates of economic value of wetland ecosystem services for flood prevention range from 18 to 7,399 2017 $US per hectare per year (Sharma et al. 2015; Ming et al., 2007; Watson et al., 2016; Zhang et al., 2017). Their value for hurricane protection services varies from 225 to 1700 2017 $US per hectare (Loerzel et al., 2017 ; Constanza et al., 2008). The consensus among natural resource scientists is that flood control is a localized benefit, and therefore harder to do a benefits transfer without first having primary data on the flood control at a particular site of interest.

G.3.0.1 Calculating a parcel–specific FRI

The predicted probabilities of flooding derived from the estimated models, are used to develop a parcel–specific flood risk index. For each parcel in the data, the FRI at t_0 is the predicted probability from the probit estimation using data from Hurricane Irma. However, the FRI at t_1 needs to be adjusted to reflect changes in impervious surface cover resulting from land development decisions in t_0. The FRI for parcel i in t_1 is:

$$Pr(\text{Flood}_{it_1}) = Pr(\widehat{\text{Flood}}_{it_0}) + \hat{\gamma}\Delta\text{Impervious surface cover}, \tag{G.2}$$

where parameter $\hat{\gamma}$ is the average estimated marginal effect of a percentage increase in impervious surface cover on the probability that a parcel floods,[42]and the change in impervious surface cover (ΔImpervious surface cover) is calculated as the difference between the original and the updated share of impervious surfaces in the area surrounding a newly developed parcel.[43] According to the estimations, $\hat{\gamma}$ is 16.8, so that a one percent increase in high intensity development covers leads to an 16.8 percent average increase in the probability of that a parcel floods during a storm.

G.4 Summary statistics and description of the generated FRI

Table G.2 presents summary statistics for the FRI variable at two different points in time. These estimations assume there are not changes to current patterns of land use (i.e., the correspond to predictions following a “business as usual”trajectory). As shown by descriptive statistics there is little difference across time periods.

In general, the majority of parcels in the sample are associated with low values of FRI and a few parcels have high values. This observation is also shown in figure G.5, which is a heat map showing the spatial variation of the FRI_0 variable. As shown in figure, parcels with the highest FRI are neighboring rivers and estuaries, namely, the Wadmalaw and Stono Rivers in Charleston, the Ashley and Cooper Rivers in Berkeley, and the Winyah Bay in Georgetown.[44] Interestingly

[42] Taking an average of the available forecasts rather than relying on just one has been found to reduce forecast error (typically measured by the root–mean squared error), often by about 15 to 20 percent (Silver, 2012). Thus, output from the two separate models are aggregated to derive an average effect.

[43] The updated share of surrounding impervious surface cover corresponds to the average percentage impervious surface cover for parcels that developed in the most recent 5 years (i.e., from 2011 to 2016).

[44] Winyah Bay is a coastal estuary that is the confluence of the Waccamaw River, the Pee Dee River, the Black River, and the Sampit River in Georgetown county.

very few of the parcels with high FRI are near the coast, suggesting that much of the flood risk in the region is that of fluvial flooding rather than coastal flooding.[45]

[45]The most common kinds of floods are coastal, fluvial, and pluvial floods. Coastal floods are the result of extreme tidal conditions caused by severe weather (e.g., surges). Fluvial floods occur when excessive rainfall over an extended period of time causes a river to exceed its capacity. They are the most common type. Finally, pluvial floods are caused by heavy rainfall and are independent of a water body.

Table G.1: Probit estimation of parcels flooding during Hurricanes Irma and Joaquin.

Dependent variable:	Inundated by Irma		Inundated by Joaquin+	
	Marginal effect		Marginal effect	
Intercept	2.32	***	-12.7	***
Acres	0.005	*	-0.001	
Elevation (m)	-2.37	***	-0.59	***
Dist. River (mi)	-12.3	***	4.6	***
Dist. Riverˆ2 (mi)	7.31	***	-6.49	***
Dist. Road (mi)	0.1		0.49	***
Dist. Roadˆ2 (mi)	-0.08		-0.19	***
Dist. Coastline (mi)	0.15	***	-0.77	***
Dist. Coastlineˆ2 (mi)	-0.03	***	-0.017	***
Dist. Charleston (mi)	-0.07	***	0.03	***
Parcel LC24++	-0.59		-0.22	
Parcel Forest Cover	-0.18		0.58	**

Continued on next page

Table G.1 (continued).

	Inundated by Irma		Inundated by Joaquin+	
Parcel Wetland Cover	0.24		8.55	***
Share of area w/in 0.25 mi. that is LC24++	7.14	*	26.5	**
Share of area w/in 0.25 mi. that is LC21-LC24+++	-0.04	***	0.01	***
Share of area w/in 0.25 mi. in Forest	-0.01	***	-0.05	***
Share of area w/in 0.25 mi. in Wetlands	0.05	***	-0.03	*
	Marginal effect		Marginal effect	
Buffer LC24×Acres	5.03	***	-0.05	
Buffer LC24×Dist.River	-20.2	*	48.2	***
Buffer LC24×Dist.Charl	-1.89	***	0.6	***
Data	CERA		USGS + buffer	
N Parcels	174,687		106,214	
N Flooded parcels (obs=1)	6,999 (7%)		2,871 (24.5%)	
Average water height	6.3 feet above MSL		3.45 inches	
Average predicted probability	0.04		0.03	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Continued on next page

Table G.1 (continued).

Signif. codes correspond to robust standard errors.

++ LC24 is the code in the NLCD for high intensity development cover.

+++LC21-LC24 are open space, low-, medium-, and high- intensity.

Table G.2: Summary statistics: parcel–specific predicted flood risk index (FRI).

	FRI_{2016}	FRI_{2021}	FRI_{2041}
Min	0	0	0
Mean	0.101452	0.14461	0.103084
Median	0.039775	0.08627	0.043034
Max	0.973027	0.94812	0.94812
# obs		53,219	

$FRI_t = FRI_0 + \hat{\gamma}\Delta_t$(Impervious surface cover).
$\hat{\gamma} = 16.8$.

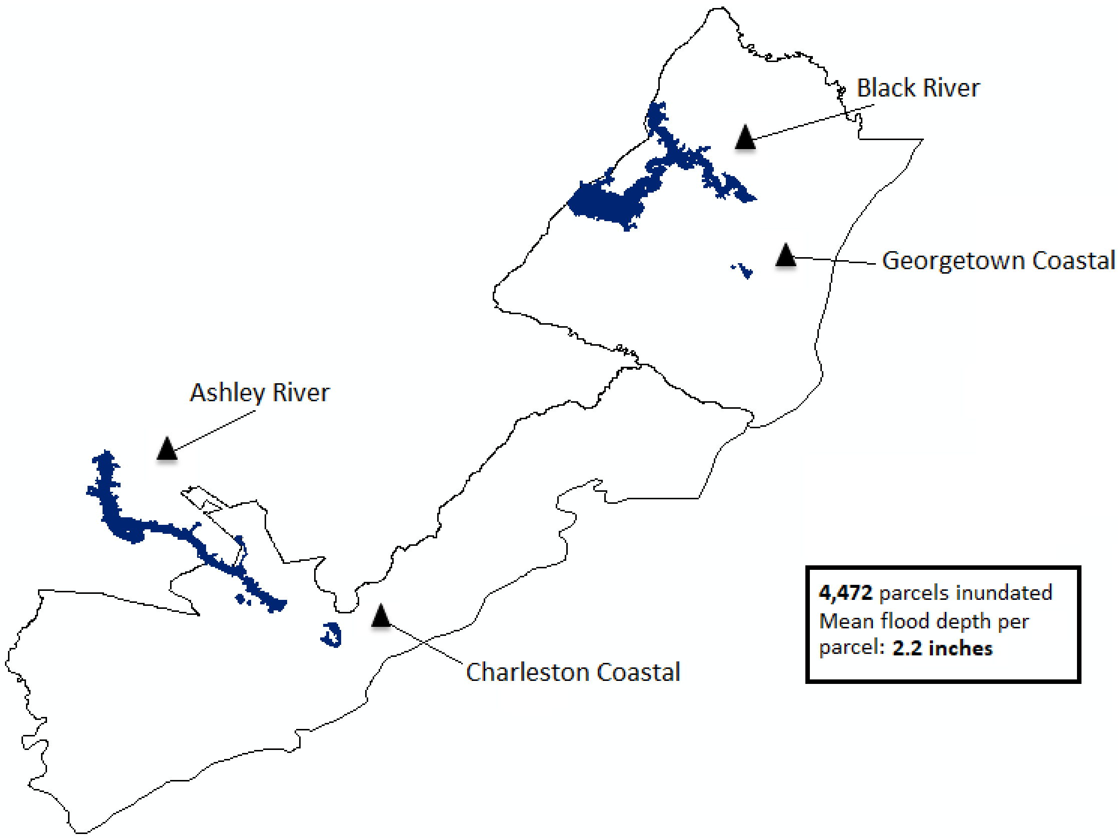


Figure G.1: Inundation maps of Hurricane Joaquin used in developing a flood risk index. Source: USGS, 2016.

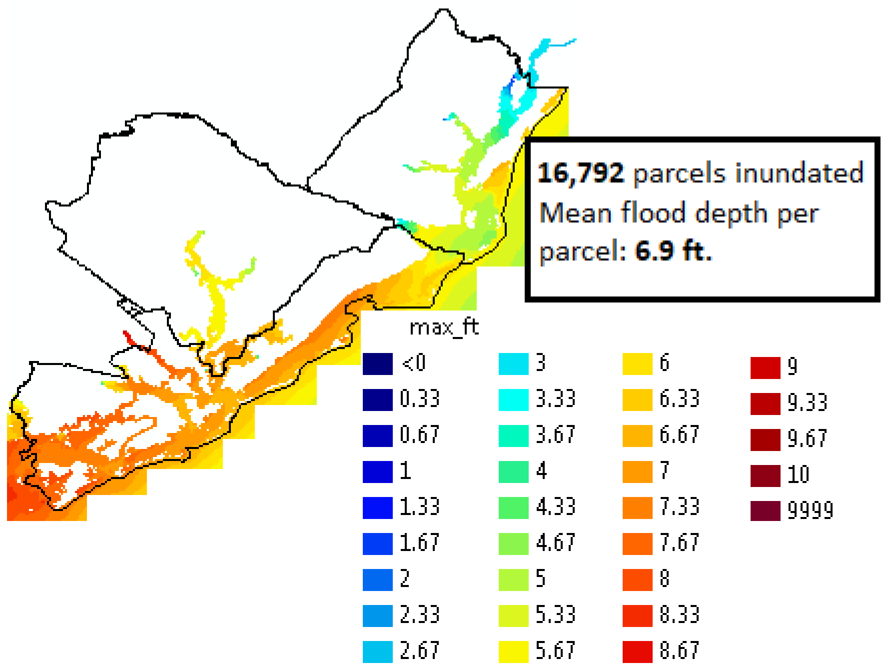


Figure G.2: Storm simulation of Hurricane Irma used in developing a flood risk index. Source: CERA, 2017.

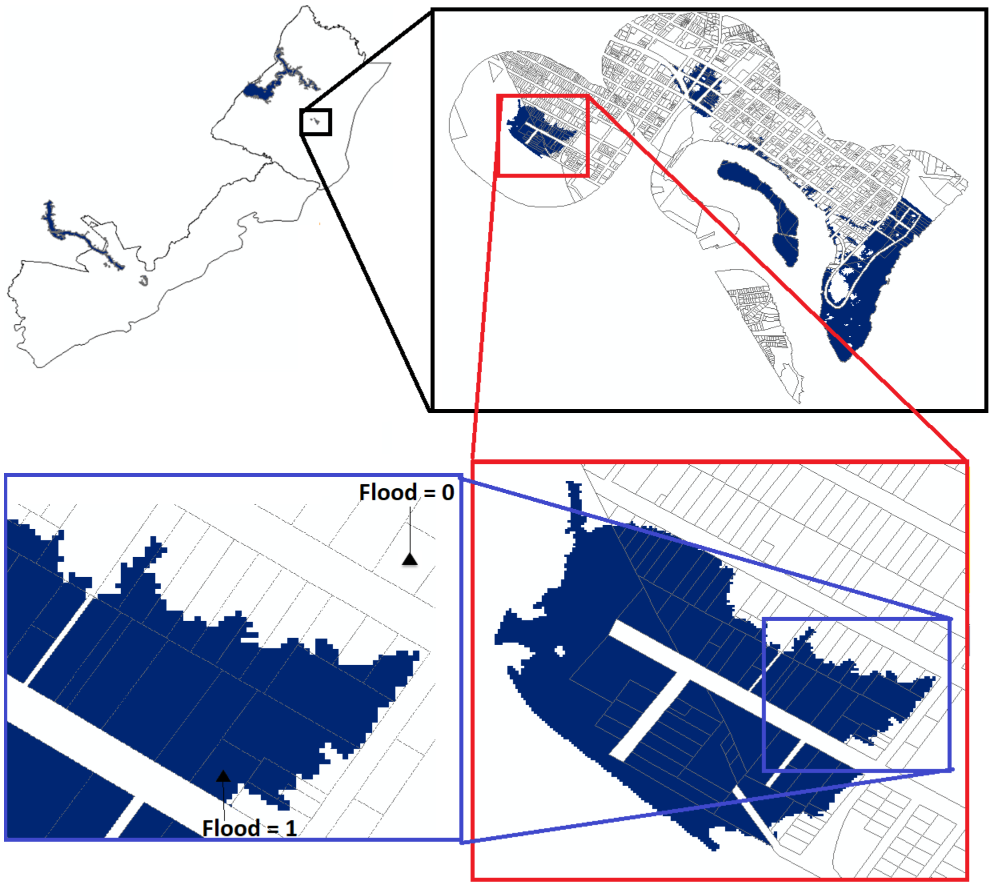


Figure G.3: Parcel selection for probit model of flood risk during Hurricane Joaquin.

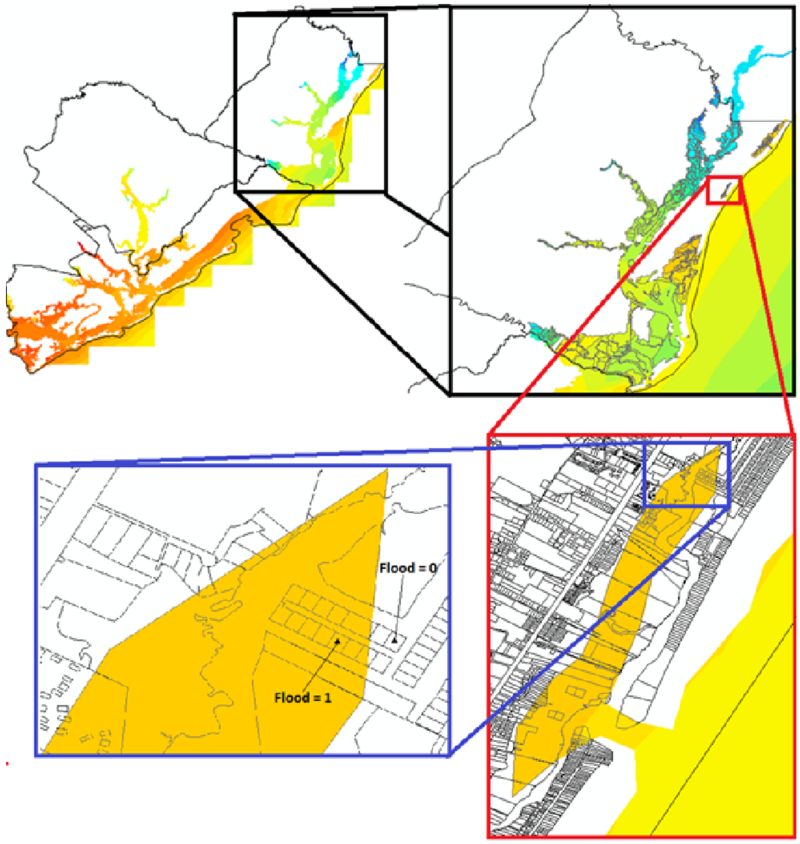


Figure G.4: Parcel selection for probit model of flood risk during Hurricane Irma.

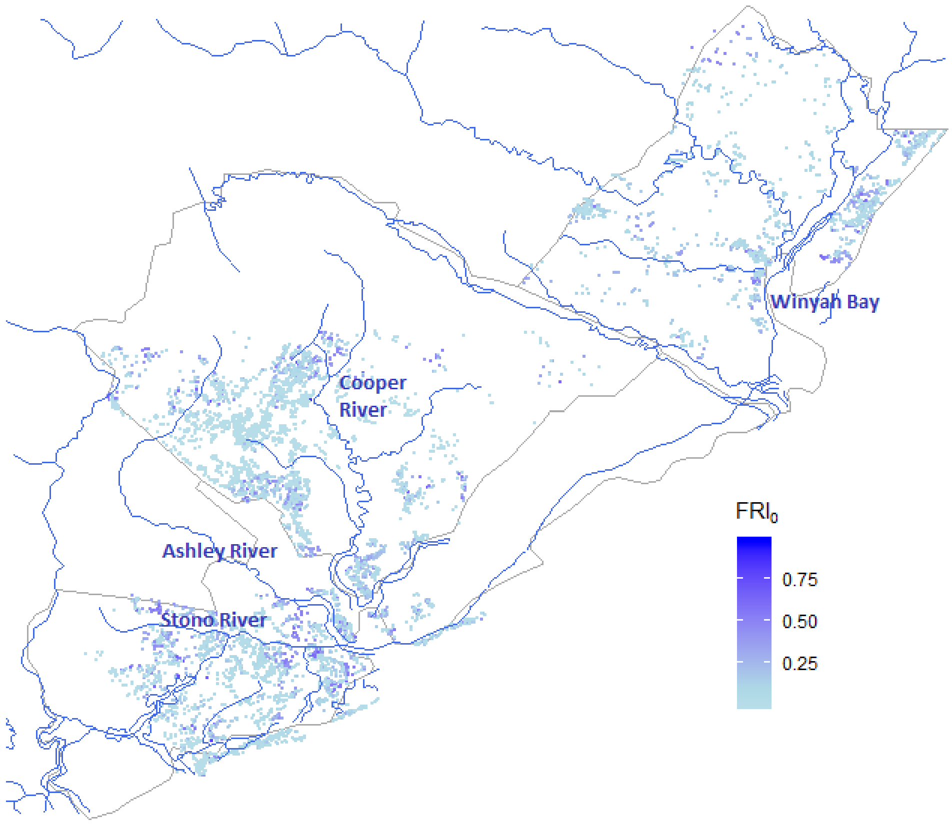


Figure G.5: Predicted FRI_0.

H Legal frame for wetland conservation

Protecting and restoring quantity and quality of coastal wetlands can play a critical role in reducing flood damages in coastal areas. In the US, the federal government derives authority to regulate wetlands from the Commerce Clause and the Clean Water Act (CWA).[46] Since 1990, when George H.W. Bush incorporated into The Clean Air Act pronouncements in Memorandum of Agreement between the U.S. Army Corps of Engineers (ACOE) and the Environmental Protection Agency (EPA) concerning the determination of the requirements for mitigating losses to wetlands, the nation's wetlands are protected by the federal government under Section 404 of the Federal Water Pollution Control Act (33 U.S.C. 1344).[47] The Memorandum established the national policy goal of no–net–loss of wetlands, a principle that enables the ACOE to ensure environmental impacts to aquatic resources are avoided or minimized as much as possible when reviewing proposed development projects.

The goal of the no–net–loss policy is to balance the protection, restoration and creation of wetlands and the need for sustained economic growth. To achieve its objectives, the no–net–loss policy establishes a flexible framework under which a variety of instruments, including tradeable wetland credits, can be utilized by federal, state and local government agencies as well as the private sector to protect and restore wetlands.

Following the 1989 directive, the ACOE and the EPA began to explicitly require property owners and developers looking to obtain a 404 permit (i.e., a permit to dredge or fill "waters of the US") to restore, establish, enhance, or preserve other aquatic resources in order to replace those unavoidably impacted by proposed development projects.[48] Under the no–net–loss policy, each newly impacted wetland is to be replaced with a surrogate wetland so that the total acreage of wetlands and the quality of wetland resources in the country does not decrease but remains constant or increases.

[46] The United States is not the only nation interested in the conservation of wetlands. International cooperation exists. The Ramsar Convention on Wetlands is an international treaty between 170 contracting parties for the conservation and sustainable use of wetlands. It is the oldest multilateral international conservation convention and the only one to deal with one habitat or ecosystem type.

[47] The first legal protection of wetlands came from President Jimmy Carter in 1977. However, it was not until the 1989 move that the government made explicit its policy goal regarding wetlands preservation.

[48] Section 404 of the CWA requires a permit before dredged or fill material may be discharged into "waters of the United States."

Despite apparent legislative clarity, even today, the level of federal involvement in wetland conservation is subject to the scope of *what* constitutes a regulated wetland (i.e., whether or not a particular wetland is part of the "waters of the United States") and that definition has changed over time.[49] For instance, two recent Supreme Court decisions (in 2001 and 2006) invalidated federal wetland regulation authority over some "isolated and non–navigable waters," and removed federal authority to protect the waters for migratory bird species. With these two rulings, the Court determined that migratory bird habitats do not constitute interstate waters, and that only wetlands that are adjacent to *navigable waters* are waters of the US and are therefore inside the scope of authority granted in the CWA. The outcome of these cases could either confirm or narrow the reach of federal regulatory jurisdiction over wetlands under the CWA.[50]

Following the 2001 and 2006 outcomes, in 2008 the US ACOE and the EPA released a new rule to clarify how to provide compensatory mitigation for unavoidable impacts to the nation's wetlands and streams. The rule defined new standards and expanded the agencies' role in regulating mitigation projects under the CWA.

More recently, in 2015, Ex–President Obama finalized the Water of the United States (WOTUS) rule, which gives the EPA and the ACOE broad authority over regulating the pollution of wetlands and tributaries that run into the nation's largest rivers. The WOTUS rule covers wetlands adjacent to either traditional navigable waters or interstate waters, as well as streams serving as tributaries to navigable waters. The rule says that wetlands and tributaries must be "relatively permanent," which means they can be intermittent. Defining wetlands this way extends federal jurisdiction to 60 percent of the water bodies in the United States. However, the relevance of this rule for future environmental policy is in question as the current administration is moving to craft a new more "industry-friendly" version.[51]

[49] With the federal Water Pollution Control Act Amendment of 1972, Congress authorized the EPA to regulate wastewater discharges as well as wetland habitat damages. The EPA's jurisdiction is limited to the "waters of the United States," which are determined as navigable or potentially navigable rivers, tidal waters, interstate wetlands, and intrastate water and wetlands that have some permanent connection or "significant nexus" with interstate commerce. Waters that lack these connections remain solely under state jurisdiction.

[50] In the 2001 Solid Waste Agency of Northern Cook County v. United States Army Corps of Engineers case, the US Supreme Court ruled that the CWA could not block Chicago area governments from turning an isolated wetland into a landfill because the wetland did not nourish interstate waterways. In the Rapanos et al v. United States (2006) case, the Court ruled to prosecute the plaintiffs, developers John A Rapanos and June Carabell, for filling 22 acres of wetland with sand in preparation for the construction of a mall, without filing for a permit. However, the Court was split over further details regarding the term "navigable waters."

[51] Having suspended the water rule, in January of 2019, the EPA is now crafting a Trump administration version, which is expected to include much looser regulatory requirements on how farmers, ranchers and real estate developers

Besides the mitigation tools established by the CWA, there are other federal policy instruments for protecting wetland habitats. These include conservation easement programs (such as the Conservation Reserve Program, the Water Bank Act, and the Wetlands Reserve Program),[52] land use programs intending to reduce wetland conversion to agricultural uses by removing incentives to produce agricultural commodities on converted wetlands (such as the Swampbuster Provision of the 1985 Farm Bill),[53] Fish and Wildlife Service conservation acts (such as the Migratory Bird Hunting and Conservation Stamp Act),[54] legal rulings (such as the Migratory Bird Rule),[55] and conditional federal grants (such as the Coastal Zone Act Reauthorization Amendments).[56]

At the state level, governments also have the authority to use various tools for addressing wetland protection, including policing powers to regulate the use of water and land (such as the administration of FEMA's building requirements), zoning authority, setting benchmarks for regulating net gain or loss of habitat, State Wetland Protection Plans, and the use of tradable wetlands credits. Additional policy instruments include private–public sector collaborations, such as educational efforts, conservation easement programs, land banking, and voluntary programs.[57] Similarly, at the local level, county and municipal governments can address wetland protection through stakeholder involvement, ordinances regarding land use, zoning and development standards, local Wetland Strategic Plans, and wetland mitigation banking tools (e.g., tradable wetland credits).[58]

must safeguard the streams and tributaries that flow through their property and into larger bodies of water. The Obama WOTUS regulation, which would have limited the use of pollutants like chemical fertilizers that could run off into small streams, came under fierce criticism from the rural landowners.

[52]The Conservation Reserve Program is a land conservation program administered by the Farm Service Agency. In exchange for a yearly rental payment, farmers enrolled in the program agree to remove environmentally sensitive land from agricultural production and plant species that will improve environmental health and quality. The Water Bank Act authorizes the US Department of Agriculture to enter into 10–year contracts with landowners to preserve wetlands and retire adjoining agricultural lands. The Wetlands Reserve Program (valid until 2014) was a voluntary program that offered landowners the opportunity to protect, restore and enhance wetlands on their property.

[53]Under the Swampbuster Provision, farmers that plant agricultural commodities on wetlands that were converted between 1985 and 1990 are rendered inelegible for program benefits until the functions of the wetlands are repaired, unless an exemption applies.

[54]The Act requires revenues from sales of duck hunting stamps to be used for buying waterfowl production areas.

[55]Until 2001, this ruling extended the jurisdiction of the CWA to wetlands for being habitats used by migratory birds that cross state lines.

[56]This program is jointly administered by NOAA and the EPA and requires states or territories to implement solutions to non–point pollution problems in coastal water.

[57]As of 2015, twenty states (including South Carolina) had formally adopted the no–net–loss goal, and six states had adopted a net–gain–focused goal. Twenty–five states had EPA–approved Wetland Protection Plans. Twenty–three states had formal wetland permitting programs that serve as their primary regulation mechanism for protecting wetlands from dredges and fill impacts. An additional 21 states relied exclusively on provisions of Section 401 of the CWA on Water Quality Certification to provide input into the dredge and fill permitting process. Six additional states (including SC) relied on 401 Certification primarily but had some other state permitting program protecting at least some portion of the state's coastal wetlands. (Association of State Wetland Managers, 2015).

[58]According to the ACOE database, in South Carolina, there are 28 approved mitigation banks, 17 pending

H.1 Coastal management in South Carolina

The State of South Carolina has been a pioneer in addressing shoreline change and other coastal hazard management issues. In 1977, the SC Coastal Zone Management Program was established through the Coastal Tidelands and Wetlands Act. The Program authorized the SC Coastal Council (now the Department of Health and Environmental Control–Office of Ocean and Coastal Resource Management) to administer a permitting program for designated *critical areas* (i.e., coastal waters, tidelands, beaches and beach/dune systems) in the coastal zone (SC DHEC, 2017). Initially, the law provided limited beachfront jurisdiction and limited guidance for decisions on beachfront development and erosion control approaches. Then, in 1987, recognizing the threats of chronic erosion, sea level rise, increased shoreline development and a lack of comprehensive beachfront planning and management, the State supported the creation of the Blue Ribbon Committee on Beachfront Management to make recommendations for regulators and legislators about development and beach management policies.

Many of the recommendations of the Blue Ribbon Committee made their way into law through the 1988 SC Beachfront Management Act (BMA), which expanded state jurisdiction.[59] However, an examination of South Carolina's recent history shows that coastal management policies have had little practical power on restricting construction and reconstruction decisions. For instance, in 1990, South Carolina eliminated a 20–feet dead zone within which the BMA established new structures could not be built and existing structures damaged beyond repair (i.e., when there has been a loss of two thirds of the value of the structure) could not be reconstructed, in response to demands of coastal landowners who, in the aftermath of Hurricane Hugo, demanded the right to rebuild their homes. The elimination of the 20–feet dead zone constitutes a clear weakening of South Carolina's coastal regulatory power.[60]

---

mitigation banks, 3 mitigation banks and 1 in-lieu fee program that have sold all their credits.

[59] The new enacted law defined beachfront policies that limited development and established regulatory standards for construction along the coastline. The BMA also required the SC Coastal Council to develop a state–wide, long–range comprehensive beach management plan after having established a baseline from which to measure erosion setbacks. The baseline is typically placed as running parallel to the shore along the crest of the primary sand dune (the dune immediately adjacent to the ocean). The BMA included a provision for resetting the act within ten years and every five to ten years thereafter. The SC DHEC completed the most recent review for baseline and set back line positions in 2010.

[60] The 20–feet dead zone established by the BMA extended 20–feet landward of the point of furthest erosion on the previous three decades. The BMA also established a 40–feet zone, a landward area to be 40 times the average annual erosion rate and *only* applied to eroding coasts (i.e., in an area where annual erosion rate was two feet a year, the setback line would be 80 feet behind the baseline). Within the 40–year zone, new construction is limited to 5,000

The impact of South Carolina's coastal policies has also been limited by other more stringent state policies that restrict reconstruction of coastal property, like the SC Septic System policy, and by outcomes of local law case rulings, like the 1992 case *Lucas* v. *South Carolina Coastal Council*, where the Supreme Court of the United States established that the BMA constituted a regulatory taking requiring compensation as it deprived a landowner (David Lucas) of the economically viable use of his property.[61]

Having found that local governments in coastal counties of South Carolina are virtually silent on the issue of land regulation for flood mitigation purposes, this work and its results have opened up a potential discussion over the role of federal legislation on environmental outcomes in coastal communities. More precisely, as discussed in this Appendix, a careful review of the legal history regarding wetland protection and restoration programs seems to indicate there is a high degree of institutional uncertainty regarding the extent and level of federal involvement in wetland conservation. Intuitively, legislative uncertainty adds risk to the market of tradable wetland credits, the largest ecosystem services market in the US and the world. Thus, the topic of regulatory uncertainty on efficiency of markets for ecosystem services is also relevant to this work.

sq feet of heated space and rebuilt homes cannot exceed the size, lateral extent, or proximity to the ocean of those replaced.

[61] David Lucas challenged the BMA as an unconstitutional taking of his property, claiming that the Act prevented him from developing his land and rendered the property valueless. The case made it to the Supreme Court, where justices found the South Carolina Supreme Court had erred in holding that the BMA was a valid exercise of the police power of the state as the Act sought to avoid a public harm and was therefore constitutional. According to the US Supreme Court, the South Carolina Supreme Court's public harm test was unworkable because it was impossible to differentiate between regulation that prevent a harm and those that confer a benefit.